UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Genetic manipulation of unique transcription factors in *Plasmodiumfalciparum*

by

Michel Robbertse

15017037

Submitted in the partial fulfilment of the requirements for the degree
Magister Scientiae

Biochemistry

Faculty of Natural and Agricultural Sciences
Department of Biochemistry, Genetics and Microbiology

University of Pretoria
2021

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

SUBMISSION DECLARATION

I, Michel Robbertse, declare that the thesis/dissertation, which I hereby submit for the degree MSc (Biochemistry) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: Michel Robbertse. DATE: 13/12/2021

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION OF ORIGINALITY

Full names of student: Michel Robbertse

Student number: 15017037

Declaration

I understand what plagiarism is and am aware of the University's policy in this regard.

I declare that this dissertation is my own original work.

Where other people's work has been used (either from a printed source, Internet, or any other source), this has been properly acknowledged and referenced in accordance with departmental requirements.

I have not used work previously produced by another student or any other person to hand in as my own.

I have not allowed, and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

SIGNATURE OF STUDENT:.

SIGNATURE OF SUPERVISOR: _____.

Acknowledgements

I would like to thank all the people at the University of Pretoria who have helped me to complete this degree. Dr Elisha Mugo for mentoring me during my master's project and giving some sound career advice, Suzan Maboane, Meta Leshbane, and Sandra van Wyngardt, for ensuring the correct and safe functioning of the labs and equipment. Dr Jandeli Niemand for her role as co-supervisor on this project and Prof Lyn-marie Birkholtz for her supervisory role of this project.

I extend my sincerest gratitude to my social support structure including my mother and father for supporting me while I finish this degree. My sister for giving me life advice and my partner for the endless hours of support shown during my degree.

I thank the South African Research Chairs Initiative and the National Research Foundation for financial support during my master's degree.

Summary

Severe malaria is caused by the *Plasmodium falciparum* protozoan. This malaria-causing parasite has a complex life cycle including a proliferative asexual stage and a terminally differentiated gametocyte stage, which is important in transmission. The complexity of this life cycle requires strict regulation of gene expression on multiple levels.

During the asexual stages, genes are expressed in a "just-in-time" manner, which is required for the rapid proliferation and changes between the asexual life cycle stages. There is also gene expression control in the gametocyte stages and especially during the transition into the gametocyte stages. This eukaryotic parasite has multiple levels of gene expression regulation including epigenetic, transcriptional, post-transcriptional, and post-translational. Transcriptional control includes the use of unique transcription factors to determine the rate of transcription as well as other regulatory elements such as enhancers and repressors.

Two of these transcription factors, PfMyb1 and Pf3D7_0603600, were found to be expressed during early gametocyte stages. It is postulated that since a higher transcript abundance is observed in the early gametocyte stages, these transcription factors are required for gametocytogenesis. PfMyb1 has already been found to be essential during the asexual stages of the parasite where it regulates transcription and is directly involved in the regulation of certain cell cycle specific genes. Pf3D7_0603600 has, as of yet, not been studied *in vivo* but, through *in silico* studies, has been annotated as a putative transcription factor due to the presence of an ARID DNA binding domain.

To study the regulatory function of these transcription factors, genetically recombinant *P. falciparum* lines were produced. For PfMyb1 a targeted gene disruption line was generated, and it was found to be essential in asexual stages which aligns with previous literature. Subsequently, a conditional knockdown line was produced to allow for further study of the importance of PfMyb1 in the gametocyte stages. This conditional knockout line was found to have a recombinant *Pfmyb1* locus which expressed GFP. It was found to be localised to the nucleus of the parasite which is expected for a transcription factor. A partially integrated targeted gene disruption line was produced for PF3D7_0603600 and was found to be non-essential in asexual stages.

In conclusion, to combat malaria the biology of the *P. falciparum* parasite should be well studied. These transcription factors are likely important for gametocytogenesis and should be further characterised using the recombinant conditional knockdown lines.

Table of contents

List of figures

List of tables

List of abbreviations

ACT	Artemisinin combination therapy
BLAST	Basic local alignment search tool
CFU	Colony forming units
DBD	DNA binding domain
DHFR	Dihydrofolate reductase
hDHFR	Human dihydrofolate reductase
IDC	Intraerythrocytic development cycle
LB	Luria-Bertani
MRE	Myb-regulatory element
NAG	*N*-acetyl glucosamine
OD	Optical density
PBS	Phosphate buffered saline
PIC	Pre-initiation complex
PTM	Post-translational modification
RDT	Rapid diagnostic test
RE	Regulatory element
SLI	Selection linked integration
TBS	Tris buffered saline
TBS-T	Tris buffered saline Tween 20
TE	Tris EDTA
TF	Transcription factor
TGD	Targeted gene disruption
TSS	Transcription start site
WHO	World Health Organisation
WT	Wild type

1 Chapter 1: Literature Review

1.1 Malaria

Malaria is still considered a global threat even though the number of cases are decreasing. The highest incidence of malaria is found in Africa, with 218 million cases and 384 000 deaths in 2019, followed by Southeast Asia with 6.3 million cases and 9000 deaths in 2019 (1). Pregnant women and infants are the most vulnerable groups, with an estimated 61 % of deaths occurring in children under five due to the mortality risk created by anaemia and low birth weights (1).

This disease is caused by parasitic protozoan *Plasmodium spp.* which is part of the Apicomplexan family. There are six different species of *Plasmodium spp.* capable of causing malaria in humans*: Plasmodium vivax, Plasmodium ovale, Plasmodium malariae, Plasmodium knowlesi, Plasmodium falciparum* (2) and *Plasmodium cynomolgi* (3). *P. vivax* remains dormant in the liver for extended periods (4) and is the most prevalent species that causes malaria outside of Africa (5). *P. ovale* does not have a high incidence but, like *P. vivax*, remains dormant in the liver (6). *P. knowlesi* mostly infects primates but has been found to infect humans and is prevalent in Southeast Asia (7). *P. cynomolgi* has recently been found to infect humans with the first case being in Malaysia (3). *P. malariae* is present in all malaria-endemic areas in the world and has been observed to have opposing seasonal fluctuations to *P. falciparum* (8). *P. falciparum* is prevalent in Africa (9) and is associated with severe malaria and death. This parasite causes a blockage in microvascular systems which causes multiple organ dysfunction including brain (10) and kidney damage (11). This parasitic protozoan is transmitted to humans by the female *Anopheles* mosquito during a blood meal.

1.2 Current control methods

The current methods of malaria control consist of transmission and disease control. Transmission control includes prevention of the mosquito vector biting the human host, prevention of infection of the human host or prevention of reinfection of the mosquito by the host during a blood meal. Transmission of this disease is dependent on several factors such as the density of the disease in the area and exposure to a mosquito vector (12). Highly implemented and effective methods to reduce exposure of an individual to the mosquito include insecticide-treated bed nets and indoor residual spraying of walls (9,12). The classes of insecticides used in these preventative methods include pyrethroids, carbamates, organophosphates and organochlorides (9). Pyrethroids are still effective in reducing the

burden of malaria in high transmission areas even though resistance has increased (12). Low dose primaquine has been approved by the World Health Organisation (WHO) as a usable transmission blocker (9). Vaccines have been developed against the malaria-causing parasite and the currently recommended transmission blocking vaccine (RTS,S AS01) shows a 30 % prevention against severe malaria in children below 5 years of age (13). Another treatment strategy is prophylaxis to prevent infection. Drugs for prophylaxis as recommended by the South African government are mefloquine, doxycycline and atovaquone – proguanil (14).

Disease control includes the treatment of infected individuals to reduce the number of parasites in the host's blood. Examples of drugs that target the parasite in the host's blood are artemisinin combination treatments (ACT) which is more effective than monotherapies due to the decreased level of resistance (15). ACTs are combinations of an artemisinin derivative drug coupled with some other slower-acting antimalarial drug with a different mode of action (16). The malaria-causing parasite has gained resistance against many of these treatments, and while vaccines are being researched, long-term efficacy is not yet known. To ensure that malaria treatment is maintained, effective diagnosis, using microscopy or rapid diagnostic testing, (RDT) (9) is needed. The RDT tests for *Plasmodium* antigens, to determine whether *Plasmodium* parasites are present, and specific proteins, to determine which species of parasite caused the infection (17). However, even with these multiple layers of prevention, diagnostics and drug treatment, resistance is a continuous threat. Therefore, an understanding of the morphology and genealogy of the malaria-causing parasite may aid in identifying novel targets which can be exploited to eliminate the parasite at any stage during its complex life cycle, to further new treatments and decrease transmission of the disease.

1.3 Life cycle of *P. falciparum*

The *P. falciparum* life cycle occurs in two hosts, namely the female *Anopheles* mosquito vector and the human host.

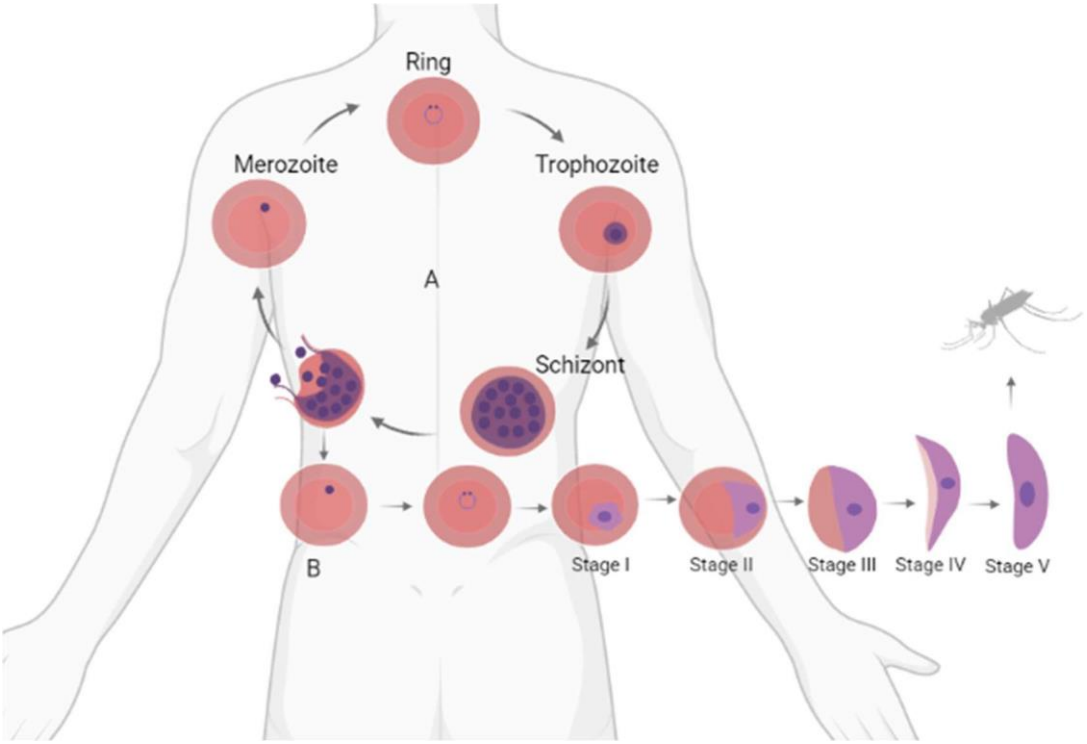

Figure 1-1. The life cycle of the *P. falciparum* parasite. The asexual and gametocyte stages occur in the human while sexual replication occurs in the mosquito. **A.** Proliferation occurs during the asexual stages with a rapid increase in the number of parasites. Invasion of the red blood cell by the merozoite occurs and it develops into a ring-stage parasite. Metabolism of the surrounding haemoglobin and nuclear division occurs. As the amount of nuclear material increases the parasite enters the trophozoite stage which is morphologically characterised as having a dense core with a "fluffy" surrounding. When the nuclear material is condensed and divides into multiple separated nuclei, it is characterised as a schizont. When the haemoglobin of the red blood cell is completely metabolised and the schizont fills the cell, the red blood cell membrane ruptures and merozoites are released. **B.** Some merozoites commit to undergo gametocytogenesis and form a sexually committed ring. Gametocytes develop in five morphologically distinct stages. Stage I gametocytes are difficult to distinguish morphologically from early trophozoites. They are circular or teardrop shaped and usually fill the erythrocyte. Stage II gametocytes form a half-circle in the erythrocyte. Stage III gametocytes start to deform the erythrocyte, elongating, starting to bow, and forming two distinct outcrops on the points. Stage IV gametocytes further elongate causing complete malformation of the erythrocyte membrane. Stage IV gametocytes are thin and pointed at both ends. Stage V gametocytes are fatter and have rounded ends. Male and female gametocytes can be differentiated during stage IV with Giemsa staining as the female has a more concentrated pigment pattern with a smaller nucleus while males have a more dispersed pigment pattern with a larger nucleus (18). This figure was created using BioRender.

The asexual stage of the life cycle occurs in humans and starts with the injection of sporozoites from the infected *Anopheles* mosquito into the human (Figure 1-1). Sporozoites then travel to the liver, invade hepatocytes, and become hepatic schizonts. The malaria-causing parasite undergoes multiple mitotic replications until the hepatocytes burst, releasing daughter merozoites into the bloodstream, where they invade erythrocytes. The parasite enters an ~48 h intraerythrocytic development cycle (IDC) where it develops into rings, trophozoites and schizonts. During the ring stage, there is a low level of transcription (19) due to the compacted nucleus and nucleosome-enriched chromosomes (20). The trophozoite stage is a growth state and thus transcription levels are increased (21). The chromatin decondenses (22) and nucleosome occupancy from the promoter regions of genes is decreased. The presence of

activating histone marks such as H3K9ac and H3K4me3 (23,24) and global transcription increases drastically (21). Once the malaria-causing parasite undergoes multiple rounds of DNA synthesis, multi-nucleated schizonts are produced. Schizogony requires a tight repacking of chromatin due to the increased mitotic replication and the subsequent release of merozoites (23). During the schizont stage, global transcription decreases (20) but there is an increase in transcription of genes involved in invasion (21). This transcription is regulated by, among others, *P. falciparum* Bromodomain protein 1, the Apicomplexan Apetela2/ethylene-1 (ApiAP2-I) transcription factor and acetylated histones (25).

Gene regulation of the merozoite stages has not been thoroughly studied, however, it can be assumed that the genome is highly compacted. Global transcription is lowest during this life cycle stage but there is an increase in transcription of invasion genes (26). The merozoites re-invade erythrocytes and the IDC starts again (Figure 1-1A).

A portion of immature schizonts commit to gametocytogenesis (27) usually due to environmental factors (28) such as low glucose concentration. These develop into sexually committed schizonts and are released as merozoites and, during the IDC, forms sexual rings (Figure 1-1B) that enter terminal differentiation. In *P. falciparum*, gametocytes mature in five morphologically distinct stages and DNA replication is decreased but metabolism is maintained. Gametocyte stages I-IV are sequestered in the bone marrow and spleen. Stage V's mature in the bone marrow before being released into the circulatory system for mosquito uptake (29). Gametocytes develop into male and female gametocytes in a ratio of one male to every four females (30).

ApiAP2-G, the master regulator of gametocytogenesis, is repressed during the IDC through enrichment of its promoter in repressive histone marks (31) and heterochromatin protein 1 (HP1) (32). *P. falciparum* gametocyte development 1 (GDV1) triggers HP1 removal from the *ap2-g* promoter which activates transcription (33). ApiAP2-G is expressed until stage II gametocytes where it is repressed by H3K36me2&3 marks on its promoter (34). AP2-G2 silences any genes required for asexual proliferation during the gametocyte stages (35). During the early gametocyte stages (stages I – III), there is an increase in heterochromatin which silences any genes not required for gametocytogenesis by an increase in H4K20me3, H3K27me3 and H3K36me2 marks (36). This correlates to the lack of DNA synthesis and nuclear division in these stages (18). Commitment to male/female gametocytes is thought to happen when sexual commitment occurs, in immature schizonts (30). ApiAP2-O3 is a regulator of female differentiation during late-stage gametocytes (37) and PfMDV1 is a regulator of male differentiation (38). The most highly expressed genes across the gametocyte stages are genes that encode for ribosomal proteins. Female gametocytes store transcripts

for later stages (39) and while these transcripts are being stored they are repressed by an RNA helicase, development of zygote inhibited (DOZI) protein (40) and the RNA binding protein, Pumilio and FBF 2 (Puf2) (41).

After ingestion by the mosquito, the gametocytes differentiate into gametes. The male undergoes three rounds of rapid DNA replication to produce 8 microgametes (38). Translational repression of the transcripts stored by the female gametocyte is removed and the transcripts are translated to form a macrogamete (39). In the mosquito midgut, the gametes come together to form a zygote that develops into a motile ookinete. The ookinete moves through the mosquito midgut and bursts, releasing sporozoites. These sporozoites then migrate to the mosquito's salivary glands to be injected into the host when taking a blood meal.

1.4 Mechanisms of regulation of gene expression in *P. falciparum*

The complex life cycle of *P. falciparum* requires stringent gene expression regulation. This tight gene expression has been observed in asexual stages with a "just-in-time" manner of gene expression (19) and a structured gene expression has also been observed in gametocyte stages (42). The "just-in-time" manner of gene expression is important for the regulation of the quick morphological changes during the IDC. Gene expression regulation occurs on multiple levels including epigenetic, transcription, post transcription and, post-translational.

1.4.1 Epigenetic regulation

Epigenetic regulation can be linked to a modification of DNA structure and has been increasingly studied in *P. falciparum* in recent years. DNA structure is either identified as heterochromatin or euchromatin. Euchromatin is classified as "open" DNA which is not tightly wound around the nucleosomes and heterochromatin is "closed" DNA which is tightly packaged around the nucleosomes. The nucleosomes consist of histone subunits and variants exist, namely H2A, H2B, H3 and H4. Modification of the 5' tails allows opening and closing of the DNA around the histone through the addition of certain groups, with acetylation, phosphorylation and methylation being the major marks. H3K9me3 and H3K36me3 are associated with heterochromatin (43) and typically decreased transcriptional activity, and H3K9ac and H3K4me3 are associated with euchromatin (44) and increased transcription. These marks are dynamic and are important in regulating DNA structure for gene expression. An example of these is the increase in H3K9ac and H3K4me3 marks during the trophozoite stages of the parasite as well as the repression of Ap2-G expression by H3K36me2&3 marks during Stage II and III of gametocytogenesis (45).

1.4.2 Post-translational regulation

Post-translational regulation includes modification of proteins which may affect their function (46). Protein modifications in *P. falciparum* include phosphorylation, acetylation and methylation, to name a few (46). Phosphorylation is a reversible process that regulates multiple processes including transcription (47) while acetylation and methylation post-translational modifications are well characterised due to their effect on the epigenome by activating or repressing gene expression (48).

1.4.3 Post-transcriptional regulation

Post-transcriptional modification includes alternative splicing, mRNA modification, mRNA decay and translational repression. mRNA modification includes the addition of a 5' 7-methylguanosine cap, a poly-A tail and alternative splicing. Alternative splicing allows differential expression of RNA transcripts through the removal of introns and joining of exons. This produces multiple different mature mRNAs from a single gene (49). Translational repression is translation inhibition of mature mRNA transcripts by e.g. RNA binding proteins (50). mRNA decay can occur from either the 5' or 3' end. Decay from the 5' end occurs when the 7-methylguanosine cap is removed and degradation of the RNA molecule by ribonucleases occur (51). Decay from the 3' end occurs when the poly-A tail is removed and the RNA molecule is degraded by the exosome complex (52).

1.4.4 Transcriptional regulation

RNA abundance is regulated through transcription levels and RNA decay. Transcription levels refer to the increased or decreased abundance of mRNA of a gene. This is regulated through a multitude of factors including *trans*-acting factors, which are proteins that directly bind DNA such as TF and *cis*-acting factors, which are DNA regulatory elements around the gene sequence, such as the transcription start site (TSS), initiator and downstream promoter elements.

Transcriptional regulation occurs when RNA polymerase II (RNA pol II) recruitment is regulated by the interplay between *cis*-acting elements and *trans*-acting elements.

RNA pol II pausing is a mechanism of transcriptional regulation where RNA pol II inhabits the upstream elements of certain genes without transcribing, thus gene transcripts can be produced quickly when required resulting in a model of all-at-once transcription. In this model, a large array of specific TF and corresponding motifs are not required, which may explain why *Plasmodium spp*. do not have a larger set of specific TFs (21).

1.5 Transcription

1.5.1 Transcription in *P. falciparum*

Transcription consists of five stages, pre-initiation, initiation, elongation, termination, and post-transcriptional processing. To initiate transcription, promoters are required to be "open" and not in a compacted DNA state (Figure 1-2A). Active promoters show a nucleosome depleted region upstream of the TSS which allows euchromatin formation and assembly of the pre-initiation complex (PIC) on the TSS (23).

In *P. falciparum*, pre-initiation starts with the assembly of the PIC. The PIC is made up of RNA pol II subunits and TFII subunits (53). Once the active promoters are available, TFIID, with a TATA-binding protein region (TBP), binds the core promoter sequence (TATA) (Figure 1-2B). Once TBP, a TFIID subunit, is bound to the promoter sequence, TFIIA-H are sequentially recruited. As with other eukaryotes, all transcribed genes require DNA elements located upstream or downstream of the gene (Figure 1-2 C). These include the TSS, the initiator sequence (INR) and the downstream promoter element. Transcription is further regulated by *trans*-acting factors, such as transcription factors, and *cis*-acting elements, such as enhancers and repressors, work to either activate or repress gene expression. These elements regulate gene expression through the recruitment of chromatin remodelling complexes as well as the PIC. The habitation of the core promoter sequence as well as the interplay between the *cis*- and *trans*-factors recruit RNA pol II which triggers basal transcription (53). Transcription co-factors are important to link the *trans*-acting factors to RNA pol II for transcription initiation. However, some PIC components, such as RNA pol II and TBP, have been identified upstream of activated and inactivated genes (54). Through this pausing of RNA pol II the "just-in-time manner" of expression is achieved by having everything "ready to go" when needed (21). As such, even though there is a paucity of TF in *P. falciparum,* they are still essential for gene expression regulation.

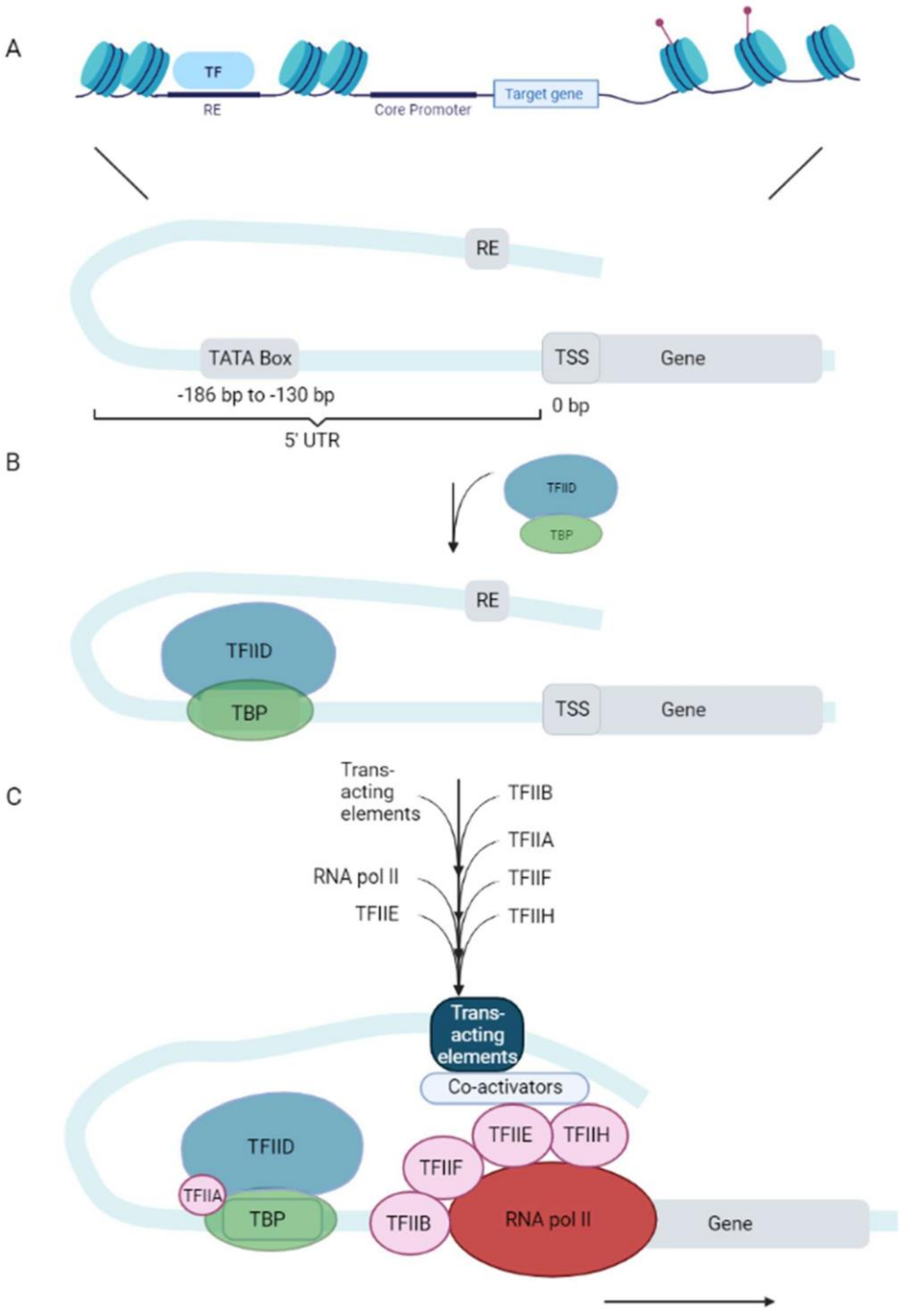

Figure 1-2. Transcription initiation in *P. falciparum*. A. During transcription pre-initiation, the nucleosomes are shifted, exposing the regulatory element (RE), the core promoter (TATA-box) and the target gene which contains the TSS. The RE are generally upstream of the core promoter and must loop around. **B.** TFIID scans the 5' UTR until the TBP domain encounters the TATA-box. **C.** *Trans*-regulatory elements and TFIIB are recruited before TFIIA-H are sequentially recruited along with RNA pol II. TFIIF functions to stabilise RNA pol II while it scans for the TSS. Co-activators are also required to bring the regulatory element in contact with RNA pol II.

1.5.2 Transcription factors

TF are proteins that are involved in transcription and usually bind DNA directly on a promoter or regulator site. Thus, TF are classified according to their specific DNA binding domain (DBD) and play an important role in promoting or inhibiting the transcription of a gene. TFs bind specific DNA motifs and recruit protein complexes required for chromatin remodelling and formation of the PIC (55). *P. falciparum* has one of the highest ratios of genes to TFs in eukaryotes (56), which suggests that TFs act as master regulators or that gene expression regulation relies heavily on other regulation methods or possibly both. TFs have been found to be essential for the asexual (57) and gametocyte stages (58) as well as the regulation of genes in response to certain environmental changes (59). There are two families of TF in *P. falciparum* that have been determined *in vitro,* ApiAP2 and Myb.

1.5.1.1 ApiAP2

The *P. falciparum* Apicomplexan AP2 (ApiAP2) family of TF is homologous to the Apetela2/ethylene response factor from *Arabidopsis thaliana* which have an AP2 DBD consisting of 60 amino acids forming a triple-stranded β-sheet stabilised by a 3' α-helix (60).

In plants, these TF either activate or repress transcription (61). In *P. falciparum* it is the largest family of TF consisting of 27 members (62) of which 21 were found in the IDC (63,64). Only two of the 27 proteins have been successfully knocked out in *P. berghei,* a species that causes malaria in rodents, showing that the rest of these proteins are likely essential in the IDC (65-68). In the asexual stages of *P. falciparum,* AP2-I was found to be essential (25), along with AP2-exp (69). ApiAP2 TF are important for the invasion process (AP2-I) as well as sporozoite formation (AP2-Sp).

ApiAP2 TF have also been found to play a role in early gametocytogenesis which suggests these TF play an important role across the entire life cycle of the parasite. AP2-G drives sexual commitment (58,69) as can be seen by its transcriptional upregulation in sexually committed rings and stage I gametocytes (Figure 1-4). AP2-G has also been found to occupy invasion genes, especially during the schizont stages (45) and is active during the first two gametocyte stages until its' expression is repressed. AP2-G2 is required for late-stage gametocyte development (beyond stage III) and suppresses any asexual proliferation genes during the early gametocyte stages (70). It has also been found to be associated with the gene necessary for male gametocyte development (*pfmdv1)* (71).

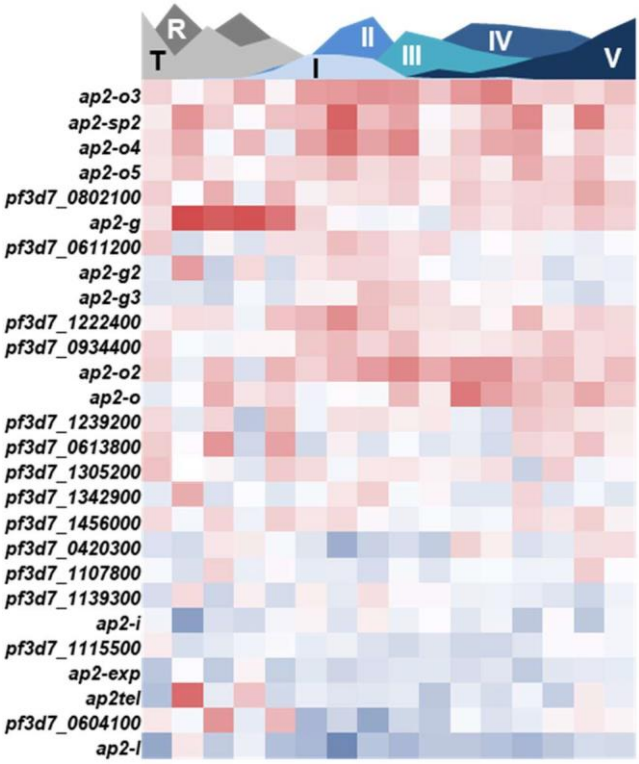

Figure 1-3. Expression levels of ApiAP2 TF in the asexual and gametocyte stages of *P. falciparum* (72). Of the 27 ApiAP2 TF, 21 are present and essential for the asexual stages of the parasite which can be seen by an increase in expression profile (red) during the asexual stages (R and T). ApiAP2-I is important for erythrocyte invasion which is affirmed by its increased expression in schizonts (here it is seen as late trophozoites) (25). However, many of the ApiAP2 TF are highly expressed throughout the life cycle with only a few required to drive gametocyte development. These include AP2-G and AP2-G2 which has been experimentally implicated in gametocytogenesis. AP2-G has been found to be an integral part in part in sexual commitment and development and has been named as the "master regulator" of gametocyte development (73) but is not expressed in stage II and III gametocytes. AP2-G2 is required for late-stage gametocyte development (beyond stage III) and suppresses any asexual proliferation genes during the early gametocyte stages (70). AP2-O3 has been implicated in female gametocyte development and since it has been found that gametocytes start sexually differentiating from stage II, it shows an increased expression profile in early-stage gametocytes. AP2-O is implicated in ookinete development (37) and AP2-Sp in sporozoite formation (66).

Some ApiAP2 TF also play a role in heterochromatin formation (PfAP2-HC) (74). Due to the essentiality and range of important functions of the ApiAP2 family of TF, it may seem like *P. falciparum* would not require other TF in the regulation of its life cycle. However, recently generated expression data from our lab showed two genes that were significantly upregulated during the initial stages of gametocytogenesis (72).

1.5.1.2 Myb

Myb proteins are characterised by the Myb DBD which consists of three tandem repeats (R1, R2 and R3) of regularly spaced tryptophan residues in a helix-turn-helix structure (75). The Myb family of TF in *P. falciparum* consists of two members, PfMyb1 and PfMyb2. PfMyb1 was initially characterised as having a single domain repeat (R2) but, when aligned with other protozoans' Myb, PfMyb1 was found to have two domains that were homologous to the slime mould *Dictyostelium discoideum* (57). PfMyb1 domains are situated in the C-terminus of the

gene which differs from other organisms where the domains are usually situated in the N-terminus (75). The conserved tryptophan residues are replaced by imperfect repeats of tyrosine, cytosine, and tryptophan residues. Myb proteins generally show DNA binding activity in regions called Myb regulatory elements (MRE).

PfMyb1 was the first transcription factor in *P. falciparum* that was functionally analysed and, using piggyBac transposon mutagenesis, has been found to be an essential transcription factor (76) needed for the parasite's progression into schizogony (75). Using ChIP-Seq, MRE have been identified in the promoter region of nine genes. These genes include H2, H3.3, serine/threonine protein phosphatase 7, proliferating cell nuclear antigen 2, a putative 26S protease regulatory subunit 8, phosphoglycerate kinase, calcium dependent protein kinase 2, the a- subunit of casein kinase 2, heat shock protein 70 and GBP130 protein (57). When treated with dsRNA, phosphoglycerate kinase, calcium dependent protein kinase 2, proliferating cell nuclear antigen 2, serine/threonine protein phosphatase 7, histone H2 and H3.3 and TATA-binding homologue were downregulated and protein kinase 5 was upregulated. Thus, this shows that PfMyb1 acts as an enhancer for the genes that were downregulated and acts as a repressor for protein kinase 5. Since some genes without MRE were downregulated using dsRNA, it can be postulated that PfMyb1 RNA also acts as a regulatory element.

PfMyb2 has not yet been functionally characterised but has been found to be essential in the asexual stages of the parasite (76).

1.6 Novel TF in *P. falciparum*

In this study, two TF (PfMyb1 and PF3D7_0603600) were identified through transcriptomic data to be upregulated during the early stages of gametocytogenesis (72). *pfmyb1* was found to be upregulated in Stage I and II gametocytes as well as Stage IV and Stage V. *Pf3D7_0603600* is upregulated in the first three gametocyte stages.

The *Pf3D7_0603600* gene contains an A-T rich interaction domain (ARID) and is annotated as a putative transcription factor as this DBD is present in many eukaryotic transcription factors (77). An orthologue of this gene in *P. vivax* has a homeobox domain, indicating that this orthologue is a TF.

The ARID DBD was first identified in mouse B-cells and was annotated as the Bright TF (78). It has since been identified as being conserved throughout the evolution of higher eukaryotes. Thus, the number of organisms that the ARID DBD is found in is widespread with multiple cellular functions, including chromatin remodelling, regulation of cell growth, differentiation and

development (79). *Pf3D7_0603600* has been found to be dispensable in the asexual stages of the parasite (67) but the essentiality in gametocyte stages is unknown.

Due to the increased expression in early gametocytes (72), it can be postulated that these TF act as important regulators for differentiation (Figure 1-5). The expression profiles of *pf3d7_0603600* during the asexual and gametocyte stages show that its expression is slightly upregulated during the ring stages and is maintained during trophozoite stages. *pf3d7_0603600* is more highly expressed in the early gametocyte stages than in any of the asexual stages. Interestingly, expression is downregulated during the initial period of stage I. This may be during the committed ring stage but is then upregulated during stage I/II and is continuously expressed until mid-stage III. It is slightly upregulated during end stage III and the start of stage IV. Mid stage IV, the expression is highly upregulated before being expressed less and only being maintained during stage V (Figure 1-5).

pfmyb1 expression increases during the asexual ring and early gametocyte stages. The expression is upregulated during the ring stage but is not as high during the committed ring stage which may suggest that PfMyb1 is not essential during the switch from the asexual to gametocyte stages but more likely plays a different role. It is highly upregulated during stages I and II and slightly less so during stage III until mid-stage IV where it is highly upregulated again before the expression decreases during early-stage V and then slightly increases during late-stage V.

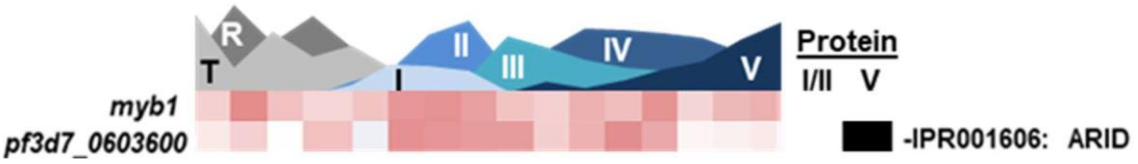

Figure 1-4. The expression profiles of *pfmyb1* and *pf3d7_0603600* genes across the life cycle. PfMyb1 is highly upregulated during the ring stages and early-stage gametocytes while Pf3D7_0603600 is downregulated in late asexual stages and highly upregulated during the first four stages of gametocytes (42). This may indicate that these two TF are important for early-stage gametocytogenesis.

It is difficult to postulate the role of these transcription factors from the expression profiles alone. Thus, the function of these transcription factors will be studied using genetic manipulation.

1.7 Genetic manipulation of *P. falciparum* parasites

Genetic manipulation of an organism allows the identification and characterisation of genes within the genome. Genetic manipulation can be done using environmental factors, e.g., generation of resistant mutants, or through transfection. Transfection is the insertion of foreign

nuclear material into an organism. The genetic material can be classified by the area where the recombinant gene will be expressed, episomal and integrative. Episomal transfection is where the foreign nuclear material is expressed outside of the organism's genome. This is usually done using a plasmid with the gene of interest under a promoter, with the gene being expressed by native machinery (80).

Integrative manipulation is the expression of the recombinant gene as part of the parasite genome (81). The plasmid does not have a promotor acting on the gene, thus it can only be expressed under an endogenous promoter of the organism. This is classified into homologous and non-homologous recombination. Homologous recombination is dependent on homologous recombination of the recombinant gene in the plasmid into the parasite genome. Non-homologous recombination is when the modified gene is not homologous to the DNA sequence of the plasmid e.g., transposon mutagenesis (82).

Genetic manipulation in *P. falciparum* is difficult. Firstly, transfection of the genetic material is not highly effective due to the multiple membranes around the parasite (83). Secondly, *Plasmodium* spp. can maintain plasmids as episomes which is not ideal in integrative manipulation (84). Thirdly, only circular plasmids are stably replicated and maintained in *P. falciparum* (85). Fourthly, homologous recombination is slow, and production of recombinant parasites take long (86). To reduce these difficulties the selection linked integration (SLI) system was used.

SLI is based on the random homologous recombination of the parasitic DNA. The advantage of this system is that parasites that contain the gene of interest episomally and parasites that have integrated the gene of interest into the genome can be selected for separately, thus reducing drug cycling time and increasing the chance of obtaining recombinant parasites. To achieve this the system uses two different drug resistance genes [neomycin and human dihydrofolate resistance (hDHRF)] for resistance against WR99210 to select for recombinant parasites. The hDHRF gene is on the plasmid backbone and allows selection for parasites that have taken up the plasmid episomally. Neomycin (G418) resistance is used to select for parasites that have integrated the gene of interest into the genome. This resistance gene is attached to the gene of interest via a skip peptide and will not be expressed unless transcribed with the gene of interest under an endogenous parasite promoter (86) (Figure 1-6).

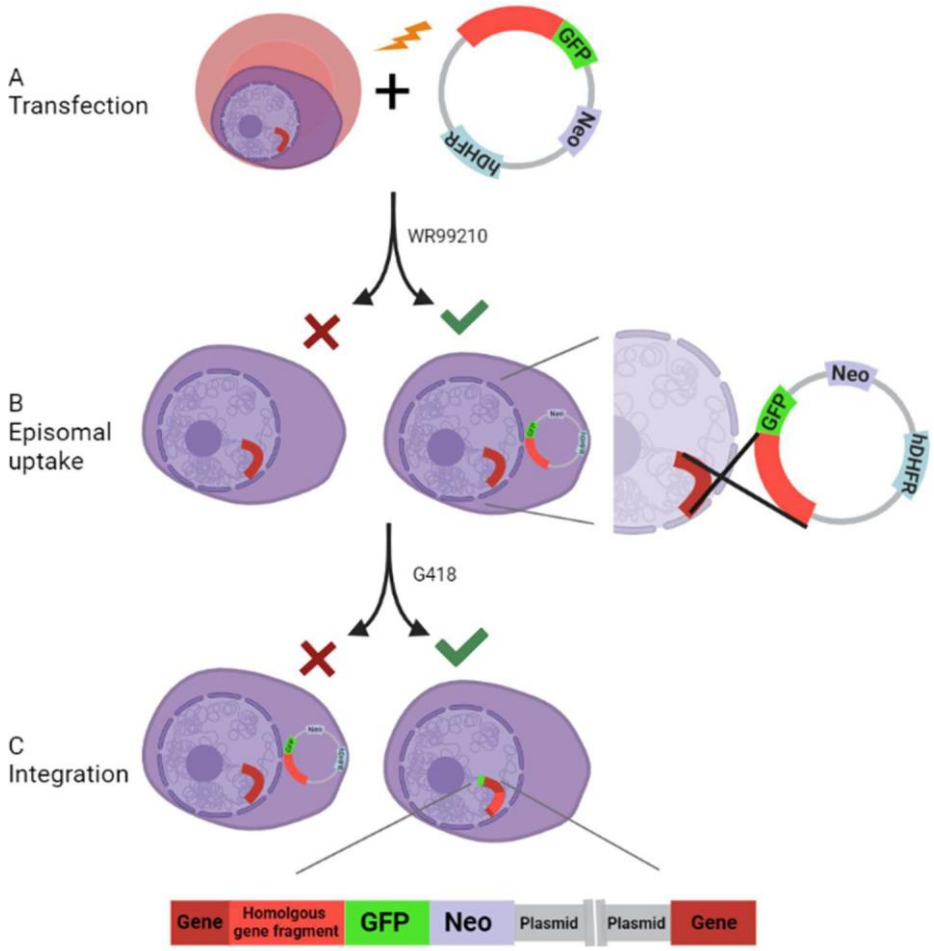

Figure 1-5. Summary of the pSLI-system used to create recombinant lines in *P. falciparum*. A. The recombinant plasmid is transfected into an iRBC using electroporation. **B.** After transfection, the parasites are treated with WR99210 to select for parasites that have taken up the plasmid episomally. Once plasmids are taken up episomally, a double homologous recombination event occurs to combine the recombinant gene into the parasite genome. **C.** G418 is used to select for parasites that have integrated the recombinant gene into the genome. The G418 resistance gene (Neo) is transcribed under an endogenous promoter so is only expressed when the recombinant gene is inserted into the genome. Created with BioRender.com.

SLI has been applied to multiple gene manipulation systems. This study used two, targeted gene disruption (TGD) and a conditional knock-down system using the *glms* ribozyme. Targeted gene disruption occurs when the gene sequence gets interrupted usually by the insertion of other DNA. In this case, it is the insertion of the pSLI-TGD plasmid (Figure 1-6). The 5' end of the gene recombines with the 5' gene fragment which was inserted into the plasmid. Due to the homologous recombination event, the entire plasmid is inserted into the genome, disrupting the gene sequence, and preventing the expression of the correct functional product (Figure 1-7). Thus, the gene sequence is modified which results in a truncated or non-functional product. This tests for essentiality i.e. whether a gene is essentially important to enable the proliferation of the parasite during the asexual stages of the life cycle.

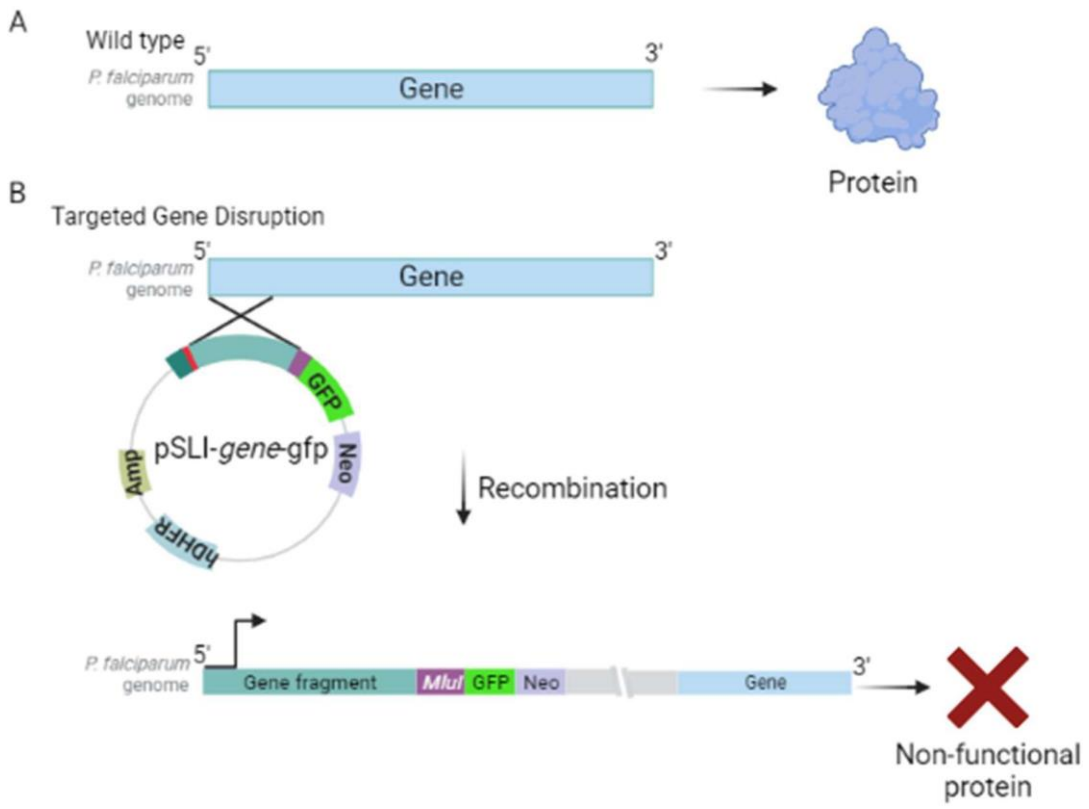

Figure 1-6. Targeted gene disruption using the pSLI-TGD plasmid. A. The wild-type gene of *P. falciparum* is not disrupted and will produce a functional protein. **B.** TGD occurs when the gene fragment, along with the pSLI-TGD plasmid, integrates into the 5' end of the gene which disrupts the coding sequence of the gene resulting in a non-functional protein.

A conditional knock-down system is the prevention of gene expression due to certain conditions. Here, the glms system was used. *Glms* codes for a ribozyme that cleaves RNA in the presence of glucosamine (Figure 1-8). There were two plasmids created for the *glms* knock-down system a WT variant and a mutant variant. The mutant variant contains a mutation on M9 to render the ribozyme ineffective. This is so that there is a comparable negative control available for when the gene of interest is knocked down.

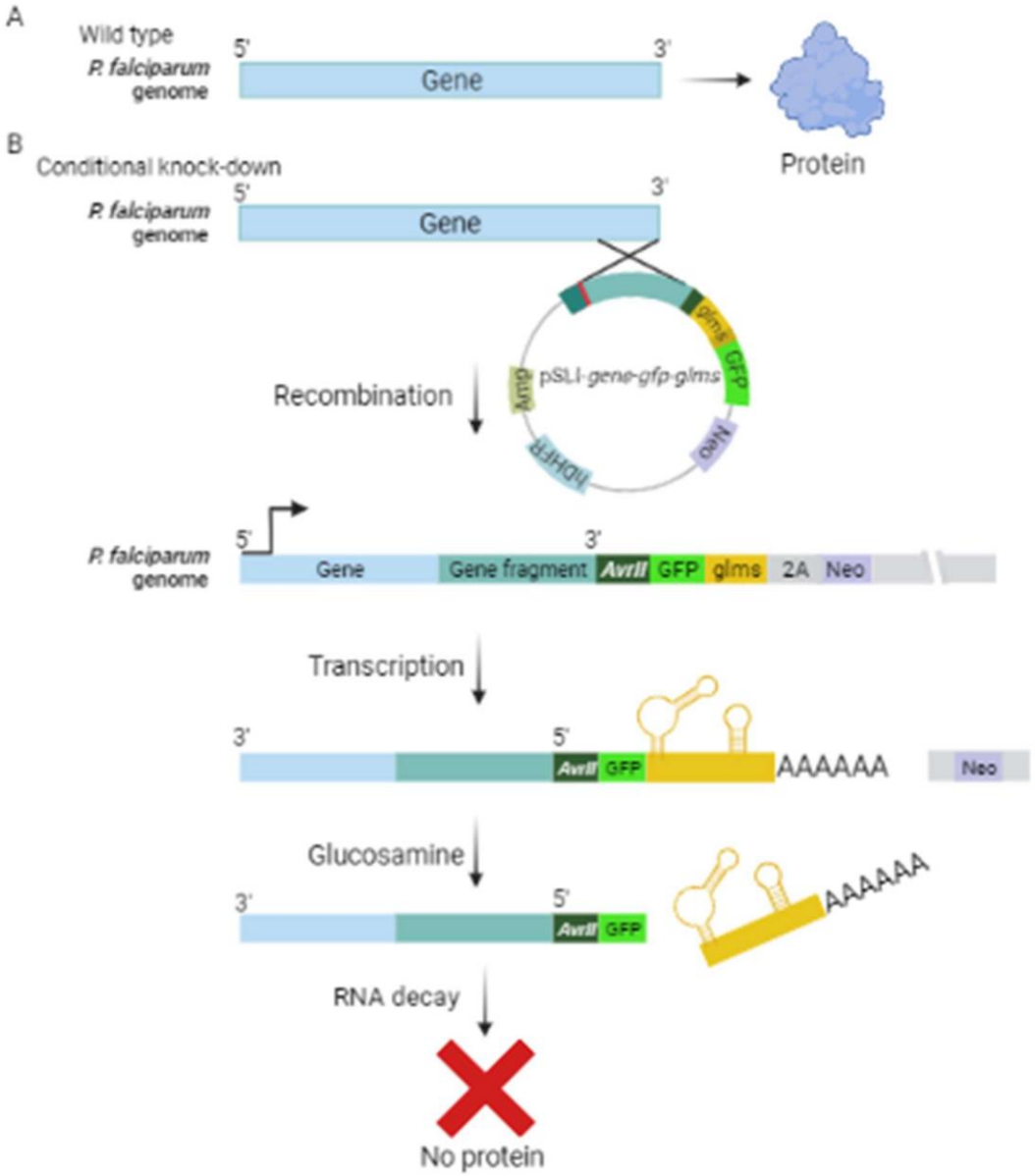

Figure 1-7. Conditional knock-down using the glms ribozyme system. A. The wildtype gene of interest will produce the expected product i.e., a protein. **B.** Once the glms ribozyme sequence is incorporated into the genome through homologous recombination, the gene will be transcribed, and the resulting mRNA will have a glms riboswitch. The addition of glucosamine will activate the riboswitch and cause cleavage which will remove the poly-A tail causing degradation of the mRNA. This results in no template for translation and no protein.

1.8 Aim

To produce genetically recombinant *P. falciparum* lines to determine the functional effects of novel transcription factors, PfMyb1 and PF3D7_0603600 during proliferation and differentiation.

1.9 Hypothesis

PfMyb1 has functional effects during proliferation and differentiation while PF3D7_0603600 has functional effects during differentiation.

1.10 Objectives

1. Produce a *P. falciparum* line to allow targeted gene disruption of *pfmyb1* to confirm essentiality in asexual stages.
2. Produce a conditional knock-down line *P. falciparum* of *pfmyb1*.
3. Produce a targeted gene disruption *P. falciparum* line of *pf3d7_0603600* to validate the essentiality in asexual stages.
4. Produce a conditional knock-down *P. falciparum* line of *pf3d7_0603600.*

1.11 Research outputs

Michel Robbertse, Riette van Biljon, Elisha Mugo, Jandeli Niemand and Lyn-Marie Birkholtz. Novel TF are important regulators in early-stage differentiation of gametocytes in *Plasmodium falciparum*. 4th South African Malaria Research Conference. Poster presentation. Pretoria, South Africa, July 2019.

Michel Robbertse, Riette van Biljon, Elisha Mugo, Jandeli Niemand and Lyn-Marie Birkholtz. Functional importance of PfMyb1 and PF3D7_0603600 TF in *Plasmodium falciparum* proliferation and differentiation**.** 6th South African Malaria Research Conference. Presentation. Pretoria, South Africa, August 2021.

2 Chapter 2: Materials and Methods

2.1 Ethics Statement

Any *in vitro* experiments pertaining to *P. falciparum* were carried out at the certified biosafety level P2 Malaria Parasite Molecular Laboratory (M^2PL). The M^2PL holds ethical clearance from the University of Pretoria's Faculty of Natural and Agricultural Sciences (reference number 180000094) for the cultivation of parasites and the Faculty of Health Sciences (reference number 506/2018) for the use of human erythrocytes. This project was approved by the University of Pretoria's ethics committee (NAS180/2019).

2.2 Bioinformatic analyses of target genes

The protein sequence for both *Pfmyb1* and *Pf3D7_0603600* was obtained from PlasmoDB (www.plasmodb.org accessed on 2 September 2020) and analysed using InterproScan (www.ebi.ac.uk/interpro/). Interpro classifies the proteins into certain families and identifies specific domains within them. This is done by determining the protein signature of the given amino acid sequence. A protein signature is a specific sequence in the primary structure of proteins that can be used to identify the domains of the protein and classify them into families (87). A protein signature is found by then to obtain protein family membership and domain classification. This is done using multiple sequence alignments that allows multiple protein sequences to be aligned simultaneously. Using a Hidden Markov Model, the amino acids are scored depending on multiple factors such as the frequency of an amino acid at a specific position. The multiple sequence alignments and hidden Markov model profile are then compared to member databases (88). This allows the annotation of the homologous superfamily, domains, and families of the protein. To get possible protein structures SWISS-MODEL repository (swissmodel.expasy.org) was used. The entire sequence is used to find a similar sequence with an existing known protein structure and the homologous regions are used to build a model of a part of the protein (89).

2.3 Production of constructs for gene recombination in *P. falciparum*

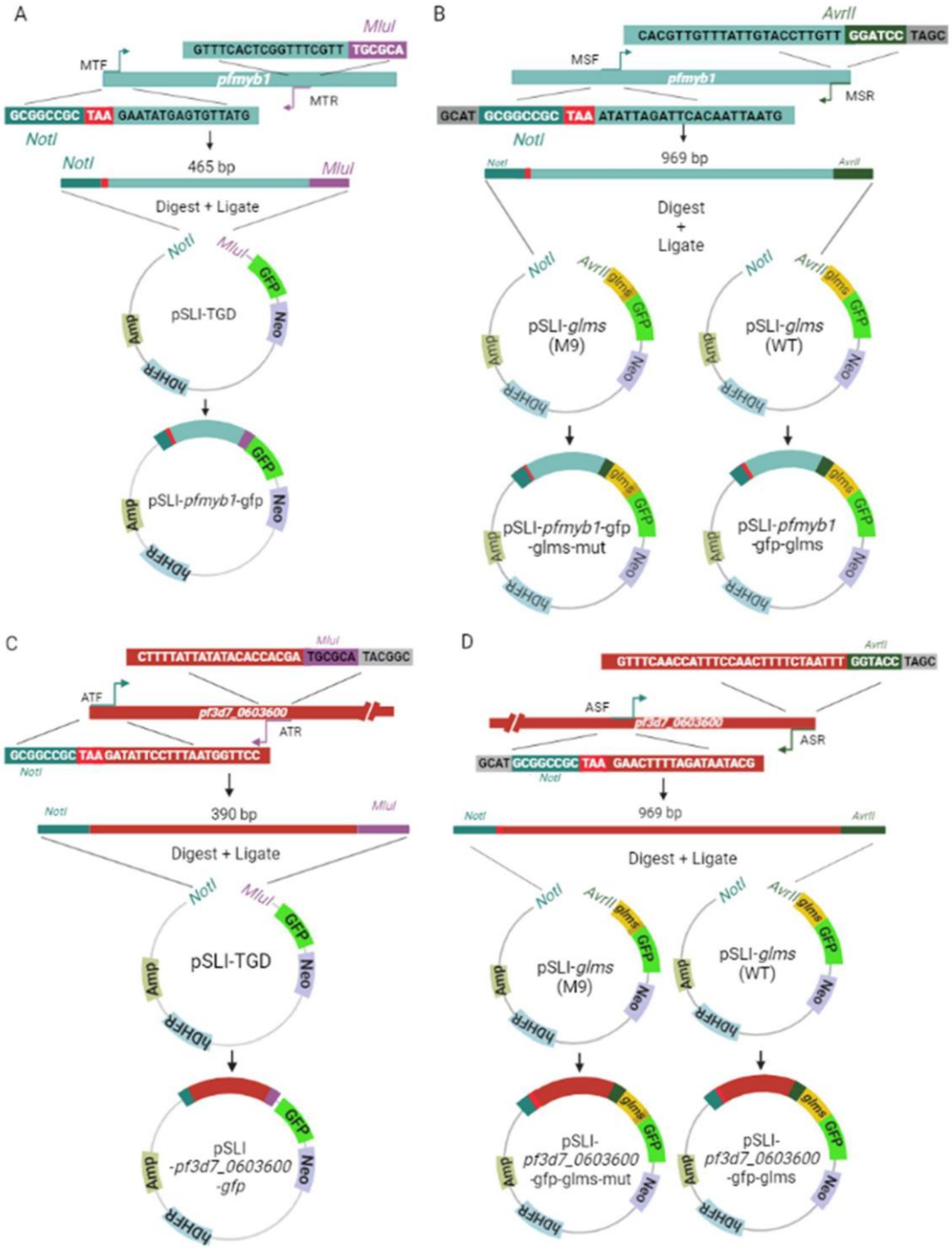

Figure 2-1. A summary of the cloning strategy of the gene fragments into specific plasmids. A. A 5' gene fragment of *Pfmyb1* was cloned into pSLI-TGD. Due to the primers used, MTF and MTR, the amplified gene fragment has *NotI* and *MluI* restriction sites. The pSLI-TGD plasmid and the gene product were digested with *NotI* *MluI.* Thus, the gene fragment has complementary sites to the pSLI-TGD plasmid allowing directional cloning. The gene fragment was ligated into the pSLI-TGD plasmid creating pSLI-*pfmyb1*-gfp which will be used for targeted gene disruption. **B.** A gene fragment from the 3' end of *pfmyb1* was amplified and the product had *NotI* and *AvrII* restriction sites due to the primers used (MSF and MSR). The product and plasmids, pSLI-*glms* (WT) and pSLI-*glms* (M9), were digested with *NotI* and *AvrII* and ligated. This created two constructs, pSLI-*pfmyb1*-gfp-glms-mut and pSLI-*pfmyb1*-gfp-glms. **C.** A 5' gene fragment of *Pf3D7_0603600* was amplified and ligated into pSLI-TGD as described in A to create pSLI-*pf3d7_0603600*-gfp. **D.** A 3' gene fragment was amplified and ligated into pSLI-*glms*

(WT) and pSLI-*glms* (M9) as described in B to create pSLI-*pf3d7_0603600*-gfp-glms-mut and pSLI-*pf3d7_0603600*-gfp-glms-mut.

2.3.1 Primer design

Primers were designed to produce a 500-1000 bp gene fragment from the 3' end of the gene with a 5' *NotI* restriction site and a 3' *AvrII* restriction site for directional cloning into pSLI-*glms* (WT) and pSLI-*glms* (M9). Additional primers were designed to produce a 300-500 bp gene fragment from the 5' end of the gene with a 5' *NotI* restriction site and a 3' *MluI* restriction site for cloning into pSLI-TGD. The parameters used when designing primers are a size of 25-35 bp, a melting temperature of 50 to 65 °C, a difference in melting temperature between primer pairs of less than 5 °C. The LiG value measures the spontaneity of the formation of a hairpin in the primer. LiG homodimer measures the spontaneity of formation of a hairpin loop with the same type of primer. LiG monomer determines the spontaneity of formation of a hairpin with itself and LiG heterodimer determines the spontaneity of formation of a hairpin structure between primer pairs. The higher the value the more likely that these structures will form, or in other words how much energy is required for the primer to bind to the template instead of itself or the other primer.

The gene sequences for *pfmyb1* and *pf3d7_0603600* were obtained from the NCBI database and exported to the online molecular biology tool Benchling (www.benchling.com). Primers were designed for each of these genes to amplify either a 5' or 3' end fragment. Primers used to amplify the 5' fragment of the gene for insertion into the pSLI-TGD vector have a 5' *NotI* restriction site and a 3' *AvrII* restriction site, the products produced are between 500 – 1000 bp (86). The restriction sites are used for directional cloning into the pSLI-*glms* (WT) and pSLI-*glms* (M9) vector. The 3' fragment for *pfmyb1* was amplified using the MSF forward primer 5' GCATGCGGCCGCTAAATATTAGATTCACAATTAATG 3' and the MSR reverse primer 5' CGATCCTAGGTTGTTCCATGTTATTTGTTGCAC 3'. The 3' fragment for *pf3d7_0603600* was amplified using the ASF forward primer 5' GCATGCGGCCGCTAAGAACTTTTAGATAATACG 3' and the ASR reverse primer 5' CGATCCTAGGTTTAATCTTTTCAACCTTTACCAACTTTG 3'. The primers for the 5' end fragment have a 5' *NotI* restriction site and a 3' *MluI* restriction site for directional cloning into the pSLI-TGD plasmid. The 5'fragment for *pfmyb1* was amplified using the MTF forward primer 5' CGATGCGGCCGCTAAGAATATGAGTGTTTATGT 3' and the MTR reverse primer 5' CGATACGCGTTTGCTTTGGCTCACTTTG 3'. The 5' fragment for *Pf3D7_0603600* was amplified using the ATF forward primer 5' CGATGCGGCCGCTAAGATATTCCTTTAATGGTTCC 3' and the ATR reverse primer 5' CGGCATACGCGTAGCACCACATATATTATTTTC 3'. All the forward primers were designed with a stop codon before the gene and after the restriction site to prevent episomal expression

(86). The primers were aligned to the *P. falciparum* genome using BLAST (NCBI, USA) to determine the specificity and *in silico* cloning was done to determine whether the GFP tag was in frame. The primers were then ordered from IDT (USA) and Inqaba Biotech (South Africa) (Table S2).

2.3.2 gDNA isolation

Genomic DNA (gDNA) was isolated from intraerythrocytic *P. falciparum* drug-sensitive (NF54) culture (5 % (v/v) haematocrit, 5-8 % parasitaemia) using a Quick DNA Blood kit (Zymo Research, USA). Genomic lysis buffer was added to packed red blood cells in a ratio of 4:1, vortexed at 2800 rpm on a CAT Ingenierburo (Germany) for 6 s and incubated at room temperature for 10 min. The lysate was added to a Zymo-Spin™ IICR Column in a 2 mL Eppendorf collection tube and centrifuged at 10 000 *xg* for one min using a Minispin® centrifuge (Eppendorf, Germany). The collection tube was discarded and replaced, and the column was washed with 200 µL of DNA pre-wash buffer and centrifuged at 10 000 *xg* for one min. The column was washed again with 500 µL DNA wash buffer and centrifuged at 10 000 *xg* for one min. The column was placed in a clean tube and 25-50 µL of DNA elution buffer was added to the column and incubated for 5 min at room temperature. The DNA was collected by centrifuging for 30 s at 10 000 *xg*.

2.3.3 DNA quantification

The DNA concentration was quantified with a Nanodrop 1000 Spectrophotometer (Thermo Fischer Scientific, USA). DNA concentration was measured at 260 nm due to the absorption of light by the aromatic rings of the purines and pyrimidines at this wavelength (90). The concentration of the DNA was measured by a conversion factor of 50 ng/µL of dsDNA per one optical density unit at 260 nm (91). Contaminants in the solvent was measured at 230 nm and 280 nm. Salt contamination was measured at 230 nm and a 260/230 nm ratio of close to 1.8 shows low salt contamination. Proteins absorb strongly at a wavelength of 280 nm due to aromatic amino acids such as tryptophan, tyrosine and phenylalanine (92). Protein contamination is determined by the 260/280 nm ratio being either above or below 1.8-2.0 (91).

2.3.4 PCR

Primers were dissolved in 1x TE buffer (10 mM tris pH 8.0, 0.1 mM EDTA pH 8.0) or ddH_2O and incubated at 37 °C for 3 h. The concentration was measured using a Nanodrop 1000 Spectrophotometer (Thermo Fischer Scientific, USA) before preparing a working solution of 10 µM.

All reactions were run in 200 µL thin-walled PCR tubes in either an Applied Biosystems 2720 Thermocycler, a GeneAmp system 9700 Thermocycler (Applied Biosystems, USA), or a

GeneAmp Thermocycler system 2400 (PerkinElmer, USA). Each reaction mix contained 1x

KAPA ReadyMix (Roche, Switzerland), 60 ng of genomic DNA from *P. falciparum,* 10 pmol of the forward and reverse primer in a final volume of 10 µL (93).

The annealing temperatures were calculated using the formula

$$Tm = 69.3\,ºC + (0.41\,GC\,\%) - \frac{650}{Primer\ length} - 5\,ºC$$

The optimal annealing temperature was found to be 52 °C. The reaction conditions for cycling were as follows: 94 °C for 3 min to denature the DNA, 30 s of denaturation at 94 °C, 30 s of annealing and 1 min of extension at 68 °C for 25 cycles (94). A final extension of 5 min at 68 °C was done to complete the extension of all fragments. Extension time was calculated by the extension rate of KAPA Taq DNA Polymerase (1 kbp/min as per manufacturer).

PCR products were visualised on an agarose/TAE gel post-stained with Ethidium Bromide (EtBr) (40 µg/mL dissolved in TAE). Agarose, along with electrophoresis allows the separation of DNA by size. A 1.5 % agarose/TAE (w/v) (SeaKem® LE agarose, Lonza Biosciences, USA), TAE [40 mM Tris-HCl (Sigma-Aldrich, USA) 20 mM glacial acetic acid (Merck, USA), 1 mM EDTA (Sigma-Aldrich, USA)] was used for separation of PCR products and a 1 % agarose/TAE was used to separate larger fragments such as plasmids. All reactions were run in a Wide Mini-Sub cell GT cell electrophoresis tank (Bio Rad, USA) at 110 V for 50 min. A 100 kbp or 1 kbp DNA molecular marker (Promega, USA) (0.5 µg) was used and either a tri-colour Blue-Orange loading dye (Promega, USA) or purple gel loading dye (NEB, USA) was used to a 1x final concentration.

PCR products were measured on a spectrophotometer to get an approximate concentration of the DNA. This is an approximate amount due to the presence of excess dNTPs and other components present in the PCR. The PCR product was digested to create sticky ends for directional cloning. All PCR products from the 3' end of the gene were digested with *NotI* and *AvrII* (10 U *NotI,* 10 U *AvrII,* 1x Cutsmart Buffer (NEB, USA), 20 µL of the PCR) at 37 °C for 3 h. All PCR products from the 5' end of the gene were digested with *NotI* and *MluI* (10 U *NotI*, 10 U *MluI*, 1x Cutsmart Buffer (NEB, USA),1 µg of PCR product). The reactions were incubated at 37 °C for 3 h and purified using a Gel and PCR clean-up kit (Macherey-Nagel, Germany). The purified DNA was measured on a spectrophotometer as described in section 2.2.3. and the total amount was found to be between 1.4-4.6 µg.

The vectors [pSLI-TGD, pSLI-*glms* (WT) and pSLI-*glms* (M9)] were isolated from bacterial cultures pSLI-TGD, pSLI-*glms* (WT) and pSLI-*glms* (M9) that were produced from plasmids that were kindly donated by Tobias Spielmann (Bernhard Nocht Institute for Tropical Medicine, Parasitology Section, Hamburg, Germany). The bacteria were grown in Luria-Bertani (LB)

liquid media [0.2 % (w/v) tryptone, 0.1 % (w/v) yeast extract, 0.2 % NaCl(w/v)] and incubated overnight at 37 °C with shaking (150 rpm). The plasmids were extracted using a NucleoSpin® plasmid extraction kit (Machery-Nagel, Germany). The bacterial culture was pelleted by centrifuging at 11 000 *xg* for 30 s and the pellet was lysed by adding buffer A1. Buffer A2 was added, the tube was inverted eight times and the reaction was incubated for 5 min at room temperature. To neutralise the lysis reaction Buffer A3 was added, and the tube inverted eight times before centrifuging for 5 min at 11 000 *xg*. The clear supernatant was loaded onto a column and centrifuged for 1 min at 11 000 *xg*. The membrane was washed by adding Buffer A4 and centrifuging at 11 000 *xg* for 1 min. The silica membrane was dried by centrifuging for 2 min at 11 000 *xg* and the DNA eluted by adding Buffer NE, incubating for 1 min at room temperature and centrifuging at 11 000 *xg* for 1 min. The vectors were digested to create sticky ends for directional cloning. pSLI-*glms* (WT) and pSLI-*glms* (M9), 1 µg each, was digested with 10 U *NotI* and 10 U *AvrII* for 3 h at 37°C. pSLI-TGD, 1 µg, was digested with 10 U *NotI* and 10 U *MluI* for 3 h at 37°C. All vectors were dephosphorylated with Shrimp alkaline phosphatase (NEB, USA). After dephosphorylation, the insert and vector backbone were separated on a 1 % TAE/agarose gel poststained with EtBr. The backbone was isolated from the gel using a Gel and PCR clean-up kit (Machery-Nagel, Germany). The DNA band was excised from the agarose/TAE gel and melted in the appropriate amount of Buffer NTI before being loaded onto the column. The sample was handled as described above.

A ligation reaction was prepared using 1x T4 DNA ligase Buffer (50 mM Tris-HCl, 10 mM MgCl2, 1 mM ATP, 10 mM DTT) (NEB, USA), 400 units T4 DNA ligase (NEB, USA), 200 ng insert and 50 ng vector as per a 1:4 ratio of vector to insert, incubated at 16 °C for 3 h and stored at 4 °C.

2.3.5 Preparation of chemically competent cells using $CaCl_2$

Competent cells are cells that can take up exogenous DNA due to holes created in the lipid membrane by $CaCl_2$ (95). A DH5a *Escherichia coli* culture was grown in LB-broth overnight at 37 °C with shaking (150 rpm) in a shaking incubator (MRC lab, Israel). The culture was then diluted 1:50 in fresh LB broth and incubated until an OD reading of 0.4, which was measured on a Gene Quant Pro (Amersham Biosciences, USA) at 600 nm, was achieved. The cells were incubated on ice for 20 min and centrifuged at 4 °C at 1865 *xg* in an Avanti J-E centrifuge (Beckman Coulter, USA). The supernatant was removed, and the pellet was resuspended in five pellet volumes of ice-cold 0.1 M $CaCl_2$ (95). The culture was centrifuged at 1865 *xg* at 4 °C for 30 min in an Avanti J-E centrifuge (Beckman-Coulter, USA). The supernatant was removed, and the cells were resuspended in five pellet volumes of a solution of 0.087 M $CaCl_2$ and 13 % (v/v) glycerol. The cells were then incubated on ice for 1 h after which these cells were used for a transformation reaction with the recombinant plasmids.

2.3.6 Transformation

Transformation is the uptake of exogenous DNA into competent cells (95). Approximately 250 ng of DNA was added to 100 µL of competent cells. The reaction was incubated on ice for 30 min, heat shocked at 42 °C for 80 s and incubated on ice for 2 min. Thereafter, the cells were recovered in 900 µL LB glucose (LB broth, 2 mM glucose) for 30 min at 37 °C in an MRC shaking incubator (MRC lab, Israel). The culture was pelleted at 1500 *xg* for 3 min. The pellet was resuspended in 100 µL LB broth and plated onto a 1 % (w/v) LB agar plate with ampicillin (100 µg/ml). The plates were incubated at 37 °C overnight.

Colony PCR was used to screen for positive colonies. A standard PCR reaction was used as described above except 1 µL of liquid bacterial culture was used as template and plasmid specific primers, SLI-F and GFP-R, were used.

2.3.7 Sanger DNA Sequencing

Positive colonies were grown up in LB broth at 37 °C overnight and plasmids were extracted using the method described in 2.3.4. The inserts were sequenced to validate correct amplification and the lack of mutations. Sanger sequencing allows detection of the base pairs using four different fluorophores which are attached to the nucleotides (96). As soon as these nucleotides are incorporated into the sequence a signal is released and recorded. The sequencing reaction was prepared using the BigDye™ Terminator 3.1 Cycle Sequencing Kit (Thermo Fischer Scientific, USA). Two reactions were prepared by adding BigDye™ Mix, 1x BigDye™ Buffer, 250 ng template and forward primer (SLI-F) 3' AGCGGATAACAATTTCACACAGGA 5' in one reaction and reverse primer (GFP-R) 5' ACAAGAATTGGGACAACTCCAGTGA 3' in the other to a final reaction volume of 20 µL. The reaction conditions were as follows: denaturation for 1 min at 96 °C, denaturation for 10 s at 96 °C, annealing for 5 s at 50 °C and extension at 60 °C for 4 min for 25 cycles. The sequencing reaction was purified using ethanol precipitation. Ethanol was added to a final concentration of 70 % (v/v) with NaOAc (83 mM) to the PCR product. The reaction was incubated on ice for 15 min and centrifuged at 10 000 *xg* on an Eppendorf Minispin® for 30 min at 4 °C. The supernatant was removed, and the pellet was washed with 250 µL of 70 % EtOH. The reaction was centrifuged at 10 000 *xg* on an Eppendorf Minispin® for 10 min at 4 °C and the supernatant removed by pipetting. The ethanol was evaporated off by incubating the reaction for 5 min at room temperature. The reaction was sequenced at the University of Pretoria's DNA Sequencing Facility on an ABI3500xl Genetic Analyzer (Thermo Fischer Scientific, USA). The sequences were analysed on CLC Main Workbench 8.1 (Qiagen Bioinformatics). Base-calling, manually assigning nucleotides to chromatogram peaks where the algorithm

was unable, was done, and a consensus sequence was produced by aligning the overlapping base pairs of the forward and reverse sequences. The consensus sequence was aligned to the *in silico* construct of the gene and vector using Benchling to determine whether there were any mutations present.

2.4 Production of genetically recombinant *P. falciparum* lines

2.4.1 Midi Prep and preparation of plasmids for transfection

Positive colonies were grown up in LB broth and the plasmid isolated using the Nucleobond Xtra Midi kit from Macherey-Nagel (Germany). DH5α *E. coli* bacterial cultures containing the plasmids of interest were grown until an OD_{600} of 3.1 to 3.8 as per the manufacturer's instructions. The cultures were pelleted by centrifuging at 4500 *xg* for 15 min (Eppendorf Minispin®) at 4 °C. Proportionate volumes of resuspension (RES), lysis (LYS) and neutralisation (NEU) buffers were calculated using the formula

$$Vol\ (ml) = \ Culture\ volume\ (ml) \text{x} OD600/50$$

resulting in culture volumes of 150 mL and 200 mL, and 6.2 mL and 16 mL of these reagents being used, respectively. The cells were resuspended in the appropriate volume of RES buffer before adding the same volume of LYS buffer and incubating at room temperature for 5 min. The lysis buffer contains sodium hydroxide and SDS to lyse the cells. Buffer NEU was added to neutralise the LYS buffer and mixed until a complete colour change from blue to colourless was observed. The filter was equilibrated with Buffer EQU before the lysate was loaded and gravity filtered. The filter was washed with 5 mL Buffer EQU and removed. The column was washed with 8 mL Buffer WASH and eluted in 5 mL Buffer ELU. This kit uses a silica-based anion exchange resin. To further increase the purity of the eluent, the DNA is precipitated using isopropanol. Room temperature isopropanol (3.5 mL) was added to the eluent and vortexed at 2800 rpm on a CAT Ingenierburo (Germany) before centrifuging at 15 000 *xg* for 30 min at 4 °C in an Avanti J-E centrifuge (Beckman-Coulter, USA). The supernatant was removed, and the pellet was washed with 2 mL of 70 % (v/v) EtOH. The mixture was centrifuged for 5 min at 15000 *xg* at 4 °C and the supernatant was removed. The pellet was left to air dry at room temperature for 10 min before resuspending in $dddH_20$. Ethanol precipitation was done by adding $1/10^{th}$ volume NaOAc and 3x the pellet volume 100 % EtOH. The mixture was resuspended and chilled at -20 °C overnight. The mixture was centrifuged at 13000 *xg* for 30 min at 4 °C in an Avanti J-E centrifuge (Beckman-Coulter, USA) and the supernatant was removed. Thereafter, 1 mL 70 % EtOH was added, and the mixture was centrifuged at 13000 *xg* for 10 min at 4 °C. The supernatant was removed and 1 mL 96 %

EtOH was added. The mixture was centrifuged for 10 min at 13000 rpm at 4 °C and most of the supernatant was removed. The mixture was centrifuged, and the rest of the supernatant was removed under sterile conditions. The DNA was resuspended in $dddH_2O$ before measuring the concentration. The DNA was precipitated again using the above method and resuspended in cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4 and 25 mM HEPES) to a final concentration of 1 µg/ul.

2.4.2 Culturing *P. falciparum* parasites

To culture *P. falciparum* parasites *in vitro* the *in vivo* conditions need to be as closely met as possible. The parasites were cultured according to the Trager and Jensen method (97-99) at 37 °C in human red blood cells (RBC) (any blood type) at approximately 5 % haematocrit. The culture media used to maintain cultures, RPMI 1640 from Sigma-Aldrich, is supplemented with 23.81 mM sodium bicarbonate, 0.024 mg/mL gentamycin, 25 mM HEPES, 0.2 mM hypoxanthine, 0.2 % glucose and 2.5 g/L Albumax II (Life Technologies, USA). The cultures are aerated with a 90 % nitrogen, 5 % oxygen and 5 % carbon dioxide gas mixture. The percentage parasitaemia is defined as the number of infected red blood cells as per the total number of red blood cells (100) and is calculated using the formula

$$Parasitaemia = \frac{no.\,of\ iRBC}{100\ total\ RBC}$$

Parasitaemia is determined by thin blood smears, fixed with methanol, and stained with Giemsa which is then viewed under an AlphaPhot 2 light microscope (Nikon, Japan) (1000x magnification, oil immersion) to determine the proportion of infected red blood cells (iRBCs).

2.4.3 Transfection

Transfection is the introduction of a recombinant plasmid to an iRBC through electroporation (101). A synchronised *P. falciparum* NF54 culture (at least 5 % parasitaemia, 80 % rings) was used for transfection. DNA (50 to 100 ug) was added to 100 µL of packed iRBCs that were resuspended in 100 µL of 1 x cytomix. The resuspended iRBCs were transferred to a 0.2 cm GenePulser®/Micropulser™ electroporation cuvette (Biorad, USA) in a Gene Pulser electroporator (Biorad, USA). The settings used for electroporation were 0.31 kV and 950 µF and maximum resistance. The transfectants were washed with media and put in culture with 5 % haematocrit in culture media as described in 2.2.8.

2.4.4 Selection of recombinant parasites for episomal uptake

Parasites were kept in a stationary incubator at 37 °C. For the first 10 days, 10 mM WR99210 (diluted in complete culture media) was added to a final concentration of 4 nM and cultured as described above. WR99210 is an antifolate drug that targets plasmodial dihydrofolate

reductase (DHFR) (102). Human DHFR (hDHFR) is a non-plasmodial enzyme responsible for producing folates for DNA synthesis that is not affected by WR99210. Thus, inclusion of the hDHFR gene confers resistance by allowing continued folate production. Thereafter, drug pressure was removed by resuspending the parasite culture in drug-free media and the transfected parasites were cultured every second day until re-emergence (approximately 1 % parasitaemia). After re-emergence the cultures were moved to incubation with shaking (~60 rpm) at 37 °C and used for DNA isolation, stocks to be frozen away and selection for integration.

2.4.5 Determination of episomal uptake

Episomal uptake was determined using PCR. DNA was isolated from the cultures following the protocol mentioned above. Primers specific to the pSLI-TGD plasmid backbone were used (SLI-F 5' AGCGGATAACAATTTCACACAGGA 3' and GFP-R 5' ACAAGAATTGGGACAACTCCAGTGA 3'). A 10 µL PCR reaction was prepared with 60 ng of DNA, 1x KAPA ReadyMix, 10 pmol each of the forward and reverse primer and dddH2O. The reaction conditions were as follows: Denaturation at 94 °C for 3 min, (denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s and extension at 68 °C for 1 min) for 25 cycles and a final extension at 68 °C for 5 min. The reaction was stored at 4 °C until visualised on a 1.5 % agarose/TAE gel as described in section 2.2.4.

2.4.6 Selection of recombinant parasites for integration

A 5 ml culture (1-2 % parasitaemia, 5 % haematocrit) of parasites that have been confirmed to have taken up the plasmid episomally was used to select for integration. The parasites were cultured as normal and G418 was added to a final concentration of 50 µg/mL daily for the first ten days. G418 inhibits protein synthesis (103). Neomycin phosphotransferase II confers resistance by enzymatic modification of the drug (104). For the next four days, G418 was added every second day and the parasites were incubated in a stationary incubator until re-emergence. The culture volume was increased for DNA isolation and to keep culture stocks.

2.5 Verification and analyses of recombination of lines

2.5.1 Determination of the integration of the recombinant locus using PCR.

DNA was isolated using the method described in 2.2.2. To determine whether wild type (WT) locus is present a primer pair was designed to flank the gene. If recombination was successful, then the product produced by the primer pair (~7.5 kbp to ~11 kbp) would exceed the extension capabilities of the KAPA Taq polymerase (3.5 kbp from genomic DNA) but if no recombination occurred then the product produced by the primer pair would be within the extension

capabilities of the Taq polymerase. The primer pair used for the locus PCR for *PfMyb1* was the PfMyb1-5' int F forward primer (5' GAAGTTTGATTGTAGTACGAATTGGTAG 3') and the PfMyb1-3' int R reverse primer (5' GTGATACATAATATCCAATATGCTAAC 3'). The primer pair used to amplify the locus region for *Pf3D7_0603600* was the A5' int F forward primer (5' GTACATATTTCCACGGGGTTAGG 3') and the ATGD 3'R reverse primer (5' CCATTCAAATTATTTGTCAGAGCACC 3') (Figure 2-2, Table S2). To determine 5' integration of *pfmyb1,* the PfMyb1-5' int F forward primer was used with the GFP-R primer. For Pf3D7_0603600 the A5' int F forward primer and the GFP-R primer were used (Figure 2-2, Table S1). For the 3' reaction, the SLI-F primer was used with the PfMyb1-3' int R reverse primer and for *pf3d7_0603600* the SLI-F and ATGD 3'R reverse primer was used. These reactions were prepared as a 10 µl PCR reaction with 60 ng of DNA, 1x KAPA ReadyMix, 10 pmol each of the forward and reverse primer and $dddH_2O$. The reaction conditions were as follows: Denaturation at 94 °C for 3 min, (denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s and extension at 68 °C for 1 min) for 25 cycles and a final extension at 68 °C for 5 min. The reaction was stored at 4 °C.

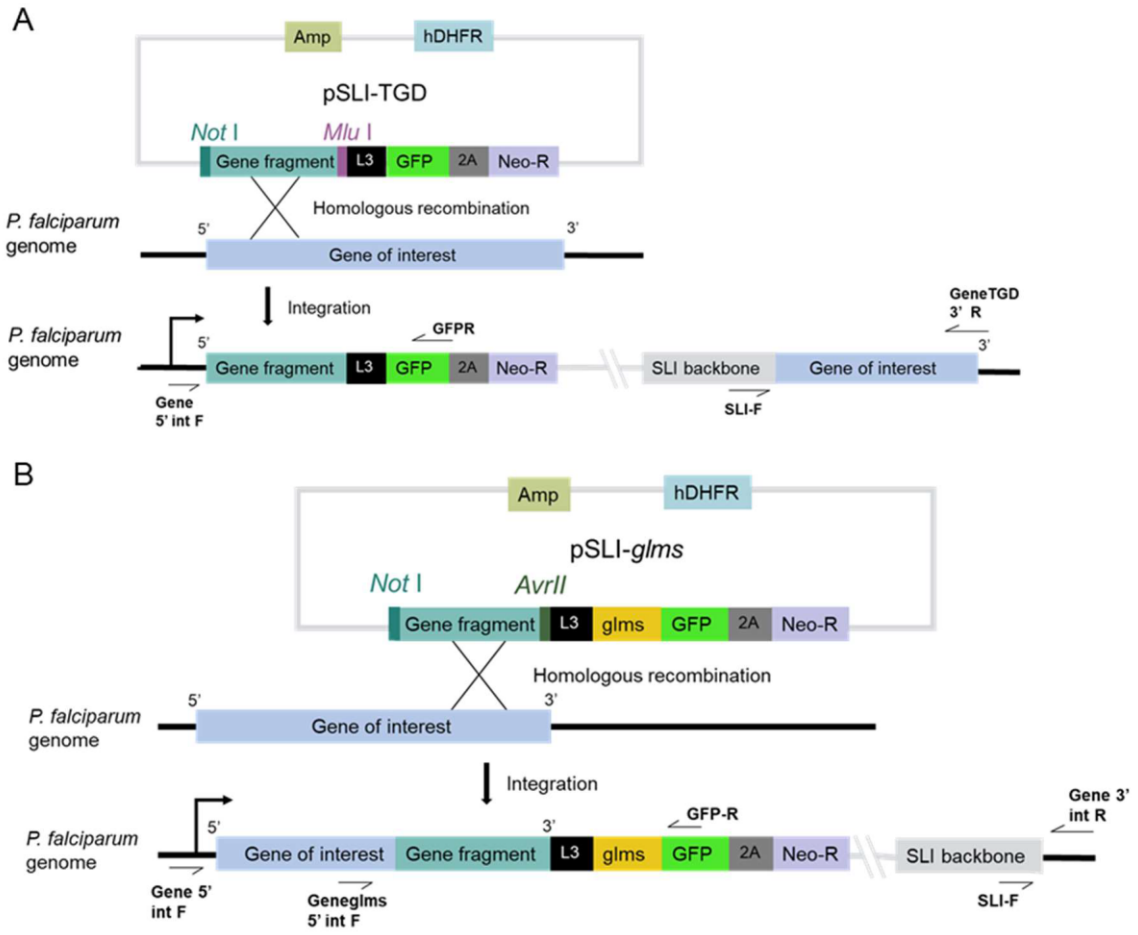

Figure 2-2. Location of primers used to test integration. A. A schematic showing the integration of the recombinant TGD gene fragment into the N-terminus of the *P. falciparum* genome. To test for the presence of the recombinant locus in the genome primers are used which span the gene fragment in the 5' upstream region and 3'

downstream region. Since the entire plasmid is recombined into the genome, a product will only be produced if no recombination occurred. To determine 5' integration the 5' primer and the GFP-R primer are used. If no recombination occurred, then GFP will not be present in the genome and no product will be produced. The same principle applies to the 3' integration PCR. SLI-F and Gene 3' int R are used and if the SLI backbone is present in the genome at the correct site, then a product will be produced. **B.** A schematic showing the integration of the 3' recombinant gene fragment into the parasite genome as well as the primers used to test for integration. The only difference is where the 5' primer is situated. Instead of the 5' UTR of the gene, it is upstream of the 3' gene fragment.

2.6 Validation of recombinant *P. falciparum* lines

2.6.1 Sorbitol synchronisation of asexual *P. falciparum* parasites

Sorbitol synchronisation was used to get the parasites to approximately a single stage. Trophozoite stage parasites from the trophozoite stage have a more permeable membrane as compared to merozoites, and rings and thus sorbitol causes the death of late-stage parasites through osmotic shock (105). An NF54 *P. falciparum* culture at 5 % haematocrit and 5 -10 % parasitaemia was pelleted by centrifuging at 2500 *xg* for 5 min at room temperature in a SL-8R centrifuge (Thermo Fischer Scientific, USA). The supernatant was removed, and the pellet was resuspended in 5 pellet volumes of 5 % sorbitol. The mixture was incubated at 37 °C for 15 min and centrifuged for 5 min at 2500 *xg*. The supernatant was removed, and the pellet was washed with 5 mL of culture media. This step was repeated 3x thereafter which the culture was put back in a culture flask, gassed and incubated at 37 °C.

2.6.2 Gametocyte production

A twice synchronised NF54 culture at 5 % haematocrit and 5-10 % parasitaemia was used for gametocyte induction. This was done by increasing the haematocrit to 6 % and decreasing parasitaemia to 0.5 % on day -3 and glucose negative media was added. Unfavourable conditions for asexual replication cause gametocyte production. On day -2 the media was not changed to allow the parasites to stress due to more due to metabolites rendering the environment unfavourable. The media was changed the next day. On the day of induction, the haematocrit was dropped to 4 %. The drop in haematocrit and glucose starvation should cause the parasites to stress and go into gametocytogenesis. The next day N-Acetyl-D-glucosamine (NAG) enriched media was added (106). NAG prevents merozoites from reinvading red bloodcells and prevents schizont maturation thus killing off asexual parasites (107). NAG media is stopped after no more asexual stage parasites are present in the culture or continued for the duration of the gametocyte cycle.

2.6.3 Dot Blot

A dot blot was done, as per the Abcam protocol, to determine the efficacy of the anti-GFP antibody. To reduce the number of interfering components the proteins were extracted into a nuclear and cytoplasmic fraction. RBC were removed by saponin lysis. A parasite culture of 8

– 10 % trophozoites was frozen at -80 °C, thawed, resuspended in a 0.1 % saponin (Merck, USA)/Phosphate buffered saline (PBS) solution [137 mM NaCl (Sigma-Aldrich), 2.7 mM KCl (Sigma-Aldrich), 10 mM phosphate (Sigma-Aldrich), pH 7.4), and incubated for 10 min at room temperature. The saponin triggers haemolysis of the erythrocyte, releasing the parasites. The saponin was aspirated, the pellet resuspended in 1x PBS and the resuspended parasite solution was transferred to a 1.5 mL Eppendorf tube. The pellet was washed until the supernatant was clear by resuspending the parasites in PBS and centrifuging at 4000 *xg* on for10 min at 4 °C. To remove the surrounding membranes, the parasite pellet was washed with double the pellet volume of Buffer 1 [0.34 M sucrose (Sigma-Aldrich), 0.5 mM Spermidine (Sigma-Aldrich), 0.15 mM Spermine (Sigma-Aldrich), 0.2 mM EDTA (Merck), 0.2 mM EGTA (Sigma-Aldrich), 15 mM Tris-HCl (pH 7.6) (Sigma-Aldrich) and phenylmethylsulphonyl fluoride (Roche)]. The cytoplasmic fraction was obtained by lysing the surrounding parasite membranes with Buffer 1 with 1 % Triton x-100 added and homogenized with a Dounce homogenizer (B pestle) (Fischer Scientific, USA) until the supernatant obtained a milky opaque appearance. The suspension was pelleted at 600 *xg* for 5 min and the cytoplasmic fraction was collected. The pellet was washed again with Buffer 1 (as described above). Half the pellet volume of low salt buffer [1.5 mM $MgCl_2$ (Merck), 0.2 mM EDTA, 20 mM HEPES pH 7.9 (Sigma), 25 % (v/v) Glycerol (Merck) and Protease inhibitor cocktail (ICN Pharmaceuticalsinc, USA)] was added dropwise to the pellet and half the pellet volume of high salt buffer was added [low salt buffer with 1.2 M KCl (Merck)]. The solution was mixed at 4 °C for 30 min andthe nuclear fraction was collected by centrifuging at 13000 *xg* for 45 min (108). This was donefor NF54-*pfmyb1*-GFP-*glms*, NF54-*pfmyb1*-GFP-*glms-mut,* NF54 and a positive control line expressing GFP.

A dot blot was done to determine whether the recombinant genes were expressing GFP. On a nitrocellulose membrane (Porablot NCP, Machery-Nagel), 2 µL of extracted protein was dotted onto the membrane in a grid formation with three spots per sample and allowed to dry. Non-specific binding sites were blocked by washing the membrane in a 5 % blocking buffer [TBS-T (20 mM Tris-HCl (Sigma-Aldrich, USA), 150 mM NaCl (Glentham Biosciences, UK), 0.05 % Tween-20 (Sigma-Aldrich, USA)] and 5 % skim milk powder (Sigma-Aldrich, USA)]. The membrane was divided into three parts: one part was incubated with polyclonal rabbit anti-GFP (Abcam, UK) diluted 1:1000, another with rabbit anti-H3 core (Abcam, UK) diluted 1:10 000 as a positive control for the nuclear fraction and the last one with rabbit anti-Hsp 70 (Abcam, UK), diluted 1:10 000 as a positive control for the cytoplasmic fraction and was incubated overnight at 4 °C. The membrane was washed in TBS-T three times for 5 min each time by adding 10 mL of TBS-T to the tube and putting it on a rotatory mixer. To visualise the primary antibody the membranes were placed in a secondary goat anti-rabbit horseradish

peroxidase (HRP)-conjugated antibody (Abcam, UK). Since the secondary antibody is conjugated to (HRP), antibody binding can be observed by chemiluminescence. The membranes were tagged with secondary antibody by incubating for 30 min at RT and washing again as described above. The antibodies were visualised using a Westpico plus Chemiluminescence substrate (Thermo Fischer Scientific, USA). The substrate was mixed by adding SuperSignal™ West Pico PLUS Stable Peroxide Solution and SuperSignal™ West Pico PLUS Luminol/Enhancer Solution in equal volumes. The solution was evenly applied to the membrane and visualised using a ChemiDoc MP Imaging System (Biorad, USA) with varying exposure times.

2.6.4 Western Blot

Whole-cell proteins were isolated from an 8 – 10 % majority trophozoite stage NF54 culture (109). The intraerythrocytic parasites were pelleted by centrifugation at 750 *xg* on a Heraeus Megafuge 40 (Thermo Fischer Scientific, USA) for 5 min at 4 °C. The parasite pellet was resuspended in lysis buffer [150 mM sodium chloride (Sigma-Aldrich, USA), 1 % (v/v) Triton X-100 (Merck, USA) and 50 mM Tris-HCl (Sigma-Aldrich, USA)]. The supernatant was removed and resuspended in 2x Laemmli Sample Buffer (69,45 mM Tris-HCl, pH 6.8, 11,1 % (v/v) glycerol, 1,1 % LDS, 0,005 % bromophenol blue and 50 mM 2-mercaptoethanol) (Biorad, USA) to a final concentration of 1x.

An SDS-Polyacrylamide gel was made to resolve the proteins. A 5 % stacking and 10 % resolving gel was used. To make the stacking gel 5 % (v/v) acrylamide mix (Sigma-Aldrich, USA), 0.125M Tris (pH 6.8) (Sigma-Aldrich, USA), 0.1 % (v/v) SDS (Merck, USA), 0.1 % (v/v) ammonium persulphate (Sigma-Aldrich, USA) and 0.1 % TEMED (Sigma-Aldrich, USA) was used. To make the resolving gel 10 % (v/v) acrylamide mix, 0.375M Tris (pH 8.8), 0.1 % SDS, 0.1 % (v/v) ammonium persulphate and 0.04 % (v/v) TEMED was used. The gel was cast using a BIORAD Mini Protean system. Proteins were loaded and the voltage was set to 60 V until past the stacking gel then 100 V to resolve the proteins. To determine the protein sizes on the gel a Dual Point Precision Plus protein ladder (Biorad, USA) was loaded as well.

To transfer the proteins to the membrane (Porablot NCP, Macherey-Nagel, Germany) the wet transfer method was used as this is the most efficient transfer method in terms of the amount of protein transferred. In the cassette of a Hoefer TE42 Tank Blotting Unit (Thermo Fischer Scientific, USA), 2 sponges, 2 filter papers, the SDS-PAGE gel, the membrane, 2 filter papers and 2 sponges were assembled. Everything was wet with transfer buffer [192 mM glycine (Merck, USA), 25 mM Tris-HCl (Sigma-Aldrich, USA), 10 % methanol (v/v) (Merck, USA)] and the tank filled. The transfer reaction was run at a constant current of 200 mA for 2 h. Once the transfer was complete the membrane was stained with a Ponceau stain solution [5 % (v/v) glacial

acetic acid, 1 g (w/v) Ponceau red] for 10 min and washed with water until bands are visible to observe the proteins on the membrane thus determining the transfer efficiency. Additionally, the SDS-polyacrylamide gel was stained with Coomassie brilliant Blue and de-stained using a de-staining solution (50 % methanol, 40 % acetic acid) to observe the proteins left behind onthe gel. The membrane was then cut before placing it in 5 % blocking buffer [TBS-T (20 mM Tris-HCl, 150 mM NaCl, 0.05 % (v/v) Tween-20), 5 % (w/v) skim milk powder) and incubated at room temperature for 1 h.

The membrane was placed in 5 % blocking buffer rabbit anti-GFP (1:500) (Abcam, UK) or rabbit anti-H3 (1:10000) (Abcam, UK) core antibody (loading control) and incubated with mixing at 4 °C overnight. The membrane was washed with TBS-T for 5 min with mixing, then twice more for 15 min. The membranes were incubated in secondary antibody [goat anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (Abcam, UK)] in 5 % blocking buffer for 1 h and washed three times before being visualised on a ChemiDoc XRS+ (Biorad, USA) using West Pico plus chemiluminescence substrate (Thermo Fischer Scientific, USA).

2.6.5 Confocal Microscopy

Confocal microscopy was used to visualise GFP as the recombinant protein contains GFP. A mixed culture (at least 8 % parasitaemia, 5 % haematocrit) with majority late trophozoites was used. The culture was pelleted, the media was washed off with 1x PBS (Centrifuged at 7000 *xg* for 1 min on an Eppendorf MiniSpin®, the supernatant was removed, and the pellet resuspended in 1 mL 1x PBS) and the haematocrit was adjusted to 2.5 % using 1x PBS. Hoechst is a blue, fluorescent dye used to stain the nuclei of cells. This dye was added to a final concentration of 1 mM and incubated for 10 min, in the dark at 37 °C. Parasites were visualised either unfixed or fixed. For the wet, the parasite suspension was washed three times as described above and 20 µL was put onto a slide, covered with a coverslip, and sealed with nail polish. For fixed cells, 40 µL of parasite suspension was put on a coverslip and allowed to dry. The cells were fixed using a formaldehyde/glutaraldehyde solution (formaldehyde 4 % (v/v), glutaraldehyde 0.025 % (v/v)) for 10 to 20 min and washed with 1x PBS. The coverslip was dried and fixed to a slide using nail polish. The nail polish was allowed to dry before viewing the slide on an LSM 880 confocal microscope (Zeiss, Germany) using Zeiss Black 2.3 SP1. Parasite structure was viewed using differential interference contrast. Hoechst has an excitation wavelength of 350 nm and an emission wavelength of 460 nm to 490 nm. GFP has an excitation wavelength of 488 nm and an emission wavelength of 500 – 510 nm. Thus, an argon laser was used at 409 – 500 nm range through a 420LP filter and thus could be detected through the blue and green channels. Images were processed in Zen Blue 3.3 (Zeiss, Germany) or Image J v 1.51 (National Institute of Health, USA).

2.6.6 Proliferation analysis of asexual parasites

Parasites from NF54, NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines were synchronised twice with sorbitol as described above and split into three cultures (0.1% parasitaemia, 5% haematocrit). Each day a slide was made, and the parasites were cultured without adjusting parasitaemia (described in 2.3.2). To accurately determine parasitaemia, 1000 erythrocytes were counted per culture per day. The parasite culture was maintained for five days. The proliferation difference of the WT and recombinant lines was tested using a two-tailed unpaired t-test with the significance cut-off at $p<0.05$.

2.6.7 Analysis of the progression of recombinant lines vs WT NF54 during the gametocyte stages

Gametocytes were produced as described above. Each parasite line (NF54, NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut*) was cultured in triplicate and the parasitaemia was determined every two days using the method described above. The difference between gametocytaemia of the WT and recombinant lines was tested using a two-tailed unpaired t-test with the significance cut-off at $p<0.05$.

3 Chapter 3: Results

3.1 PfMyb1

3.1.1 *In silico* domain classification of PfMyb1

PfMyb1 has previously been experimentally characterised as a transcription factor (57). However, to confirm this an *in silico* analysis of the predicted protein sequence was done. Firstly, the protein was characterised using Interpro, an online tool. This tool classified PfMyb1 as being part of the homeobox-like superfamily (Figure 3-1A). Members of the homeobox-like/homeodomain superfamily are transcription factors that are grouped based on the inclusion of the helix-turn-helix (HTH) DNA binding homeodomain (110). Thus, since PfMyb1 has an HTH DBD it can be classified as a transcription factor. PfMyb1 was further classified as containing a SWI3, ADA2, N-CoR and TFIIIB (SANT) DBD. These DBDs are involved in transcriptional activation and initiation, and chromatin remodelling. The SANT domain is classified as a protein-protein interaction module. It was further found to have a MYB-like DBD, thus verifying it to be a Myb transcription factor. The Myb-like domain can be involved in either protein-protein interactions, due to the similarity shared by the SANT domain, or in DNA binding due to the Myb domain. This DBD consists of 11 residues. The structure of PfMyb1 was determined using a homology model approach. Using the SWISS-Model repository, three possible structures were obtained. The structure that was chosen was found to share a 22.4 % sequence identity with the Reb1-Ter complex of *Schizosaccharomyces pombe* and showed the Myb DBD. R1, R2 and R3 can be seen in the structure (Figure 3-1B), further verifying it to be a Myb TF.

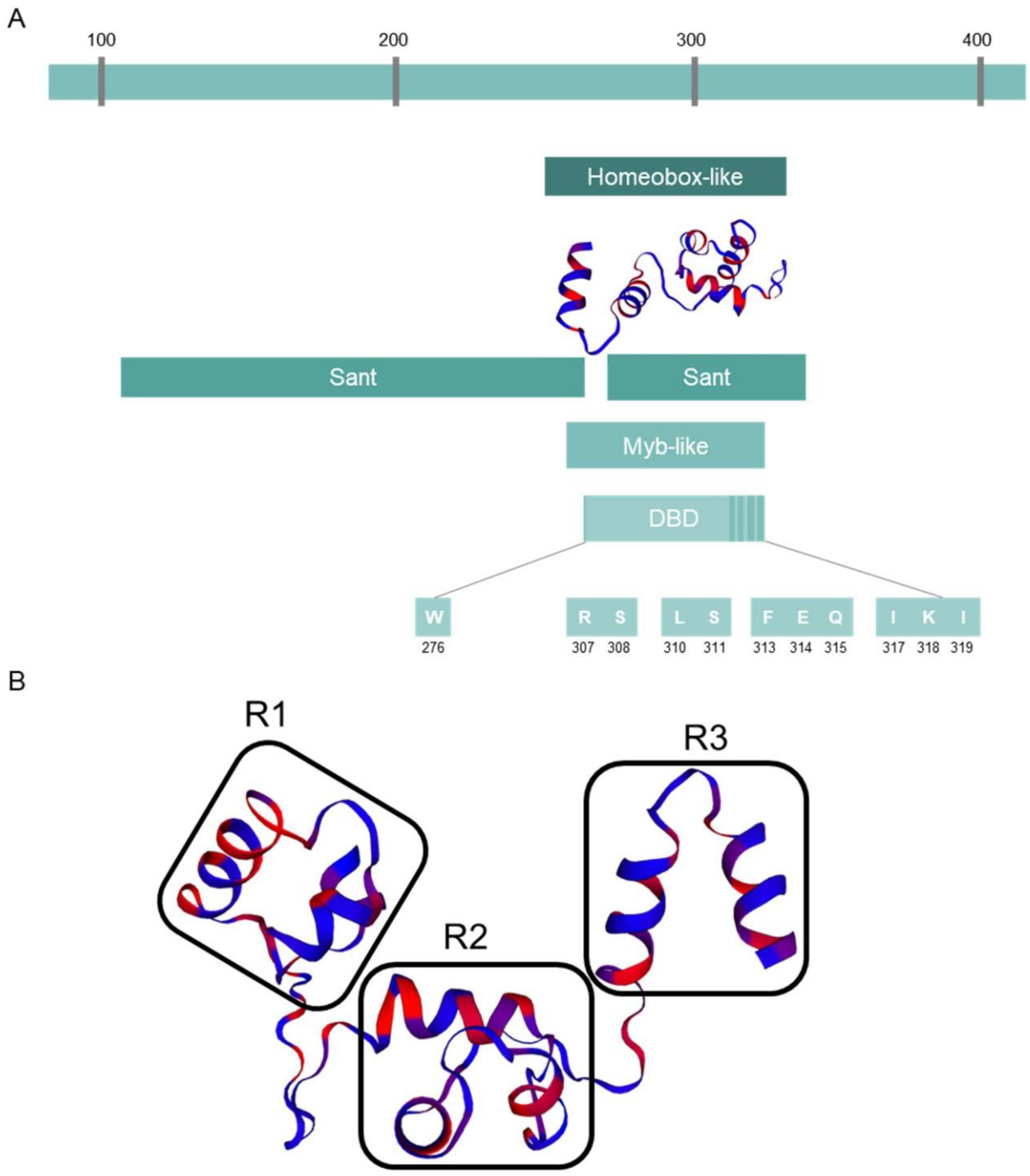

Figure 3-1. *In silico* analysis of PfMyb1 to verify its identity as a TF. A. Classification of the PfMyb1 protein. It was classified as being part of the homeobox-like domain which verifies it as a TF. It was further verified as being a Myb TF by the presence of a Myb-like DBD. A SANT domain was also present suggesting that PfMyb1 takes part in protein-protein interactions. **B.** A homology model was used to determine the structure of PfMyb1. The largest coverage available was 22 % but this included the Myb-like domain. There are three Myb domains present, R1, R2 and R3 further verifying it to be a Myb TF.

TGD requires the disruption of the functional domain of the protein. Thus, it is important to annotate the domains of the protein to ensure the disruption of the gene. Bioinformatic analysis shows that PfMyb1 contains a SANT/MYB domain, a DBD domain and is part of the homeobox superfamily (Figure 3-1).

3.1.2 Design of gene-specific primers to amplify the 5' end of *pfmyb1*

To determine the essentiality of PfMyb1, a PfMyb1 TGD parasite line had to be produced. This was done by the insertion of a 5' gene fragment of *pfmyb1* into the pSLI-TGD plasmid. The 5' gene fragment was produced using two primers, MTF and MTR. The MTF and MTR primers are 32 and 28 bp (Table 3-1), respectively. The melting temperatures (Tm) were within 0.1 °C of each other with MTF having a melting temperature of 63.1 °C and 63.0 °C. The LiG homodimer for MTF is -15.4 kcal and MTR is -10.2 kcal. The LiG monomer for MTF was 0 kcal and -0.6 kcal for MTR and the LiG heterodimer was -6 kcal. The primers should produce a product size of 490 bp. The primers were also aligned to the whole *P. falciparum* genome to determine binding specificity using Primer-BLAST (NCBI). The primers were found to be specific and to produce only a product of 465 bp of *P. falciparum* sequence.

Table 3-1. Properties of primers used to amplify the 5' gene fragment of *pfmyb1.*

Name	Size (bp)	Tm (°C)	LiG homodimer @ 37 °C (kcal)	LiG monomer @ 37 °C (kcal)	LiG heterodimer @ 37 °C (kcal)	Product size (bp)	LiTm (°C)
MTF	32	63.1	-15.4	0	-6	490	0.1
MTR	28	63.0	-10.2	-0.6			

3.1.3 Cloning *pfmyb1* into the pSLI-TGD system

The required 5' gene fragment was produced using the MTF forward and MTR reverse primers. The PCR yielded a thick band of approximately 450-500 bp which corresponded to the expected size of 490 bp (Figure 3-2A). The PCR products were digested with *NotI* and *MluI* to form "sticky-ends" that correspond with those of the pSLI-TGD plasmid. Next the "sticky-ends"of the pSLI-TGD plasmid were produced by digesting the plasmid with *NotI* and *MluI*. The insert was separated from the vector using agarose gel electrophoresis. Two bands were produced - one of approximately 7000 bp and another of approximately 2000 bp (Figure 3-2 B). The pSLI-TGD backbone is 6763 bp in length, which corresponded to the size of the larger band. Thus, it can be deduced that the 7000 bp band is the pSLI-TGD backbone.

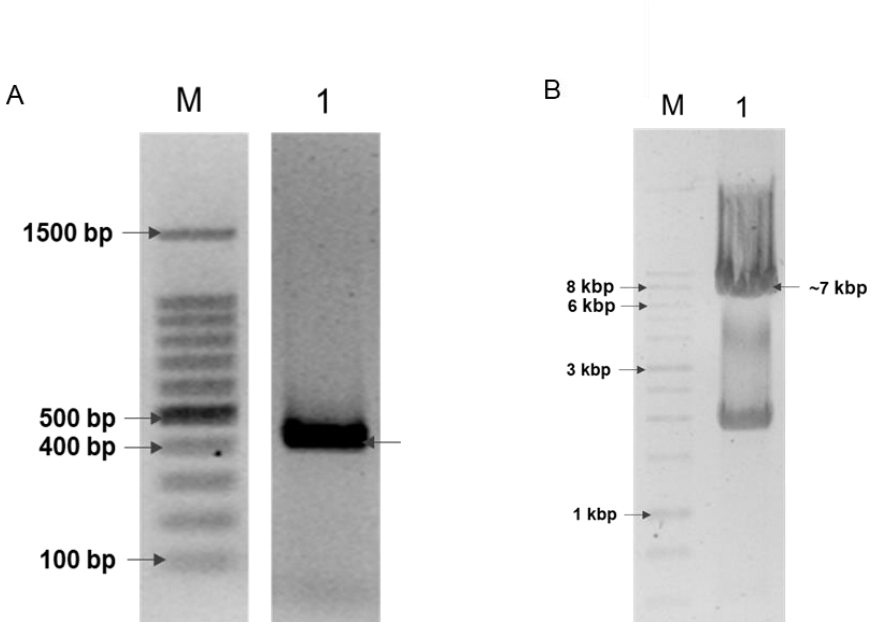

Figure 3-2. Production of the 5' gene fragment of *pfmyb1* and digestion of the pSLI-TGD plasmid. A. The *pfmyb1* 5' gene fragment was produced using the MTF and MTR primers. Lane M: 0.5 µg of a 100 bp DNA ladder. Lane 1 shows a band of approximately 450 bp was produced which corresponded to the expected size of 490 bp. The PCR was visualised on a 1.5 % agarose/TAE gel post-stained with EtBr. **B.** Restriction enzyme digestion of pSLI-TGD with *NotI and MluI* to produce sticky ends for directional cloning. Lane <: 0.5 µg of a 1 kbp ladder. Lane 1 shows two bands of ~ 7 kbp and ~ 1.8 kbp. The larger band corresponded to the expected size of the pSLI-TGD backbone (6.8 kbp). The bands were separated on a 1 % agarose/TAE gel poststained with EtBr

Since the insert had been produced and the vector and insert had been digested, the 5'*pfmyb1* gene insert was ligated into the pSLI-TGD backbone. The ligation product was transformed into DH5α *E. coli* bacteria to allow selection for colonies that have taken up the plasmid through drug selection with ampicillin. The resulting colonies were tested by PCR using the MTF and MTR primers to determine which colonies have taken up plasmids with the insert. All six colonies that were tested produced a band of approximately 500 bp (Figure 3-3A) which corresponded to the expected band size of 490 bp. Colony 2 was sequenced with the SLI-F and GFP-R primers to confirm the integrity of the insert, i.e., no mutations present and the restriction sites being intact. The chromatogram produced was clear (Figure 3-3B) showing the good quality of the sequencing reaction. Both restriction sites, the pSLI-TGD backbone and the *pfmyb1* insert were identified and no mutations were observed when the sequencing data were aligned to the *in silico* generated construct.

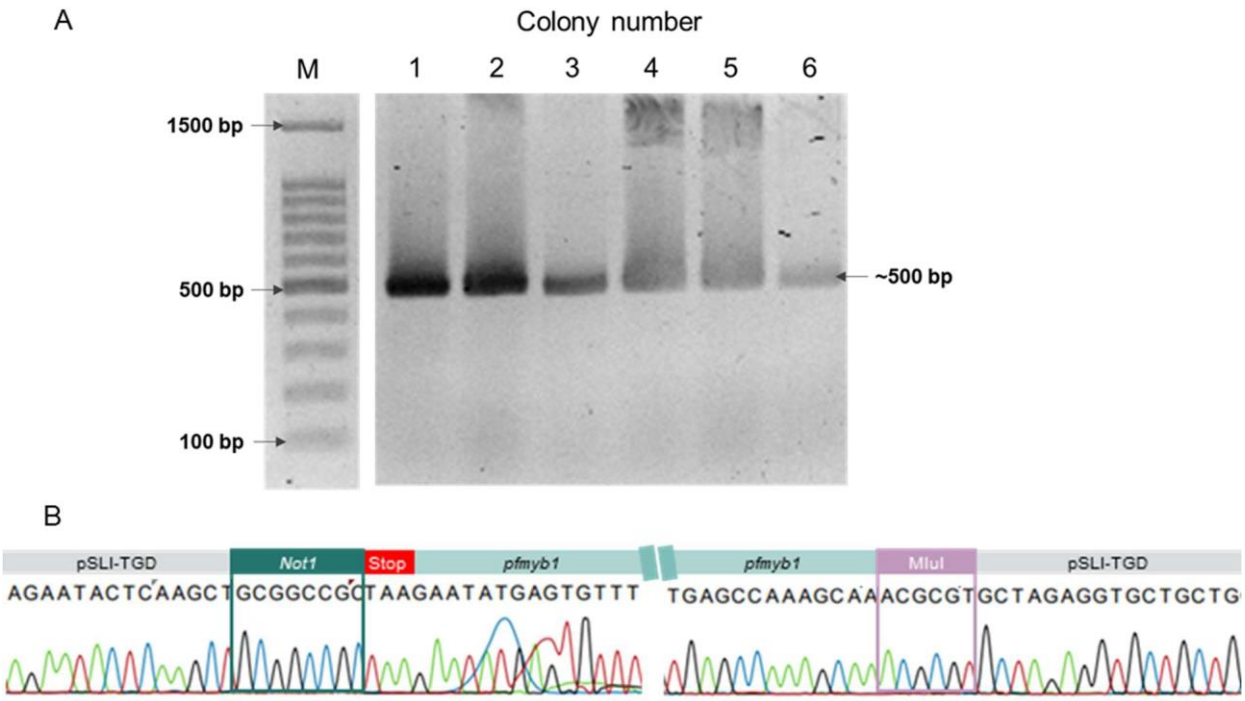

Figure 3-3. Confirmation of successfully cloned pSLI-*pfmyb1*-gfp constructs. A. A colony PCR was used to screen for bacteria positive for the pSLI-*pfmyb1*-gfp plasmid. Lane M shows 0.5 µg of a 100 bp ladder and Lanes 1-6 shows the positive colony PCR of six transformed colonies. A band of ~500 bp was observed which corresponded with the expected size of 490 bp. **B.** Colony 2 was chosen for sequencing. The restriction sites were intact, and no mutations were observed.

3.1.4 Production of NF54-epi(pSLI-*pfmyb1*-gfp).

The successfully constructed pSLI-*pfmyb1*-gfp plasmid (51.2 µg) was then transfected into *P. falciparum* NF54 majority ring-stage parasites using electroporation. The culture media was removed, and the culture was resuspended with the pSLI-*pfmyb1*-gfp plasmid. The time constant was within the accepted range of 12 ms to 15 ms (93) at 16.5 ms. The time constant gives an indication of the length of time the pores are open in the cell membrane. If it is below 12 ms then the plasmid cannot cross the membrane through the holes and if it is above 15 ms there is a high mortality rate as can be seen by the cell debris in the cultures using microscopy (111). The cultures were without drug pressure to allow the parasites to recover before starting the drug selection process. During drug selection with WR99210, there were no observable parasites after 5 days and parasites were observed after 26 days (16 days after drug pressure was removed) (Figure 3-4A). The presence of the pSLI-*pfmyb1*-gfp plasmid was confirmed through PCR using backbone specific primers, SLI-F and GFP-R, and the expected band of approximately 650 bp was produced (Figure 3-4B). These parasites were then confirmed to have taken up the plasmid episomally.

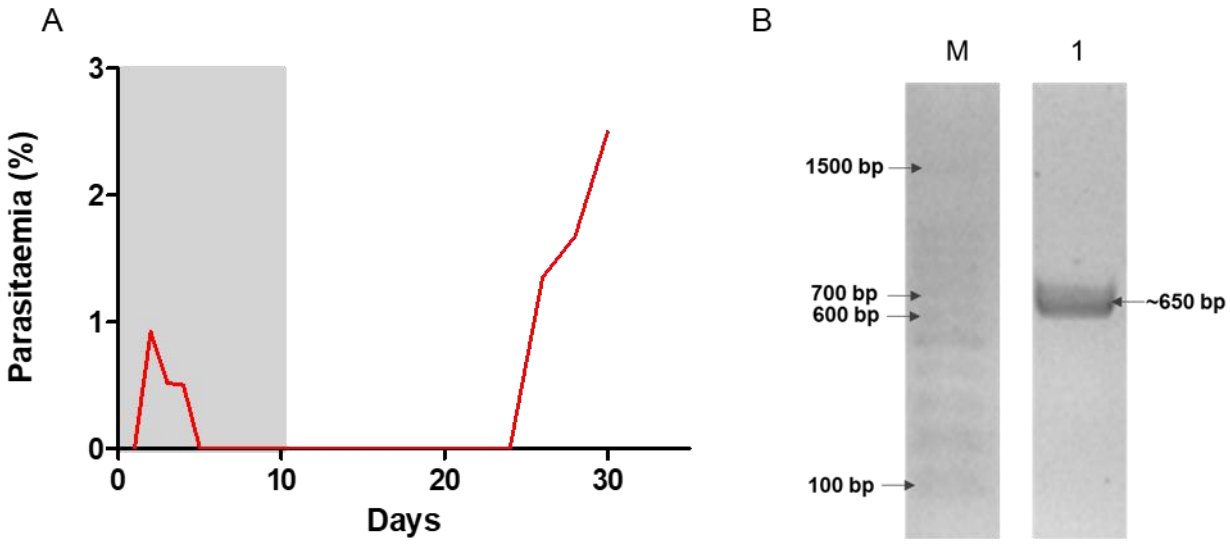

Figure 3-4. Production of NF54-epi(pSLI-*pfmyb1*-gfp) recombinant lines. A. Parasitaemia during selection for episomal uptake. The grey bar indicates that the parasites were treated with WR99210 for 10 days before allowing them to recover. **B.** Confirmation of episomal uptake by PCR. The reaction was visualised on a 1.5 % agarose/TAE gel post-stained with EtBr. Lane M-A 1 kbp ladder from Promega was used. Lane 1 shows a band of approximately 650bp.

3.1.5 Selection for NF54-Li*pfmyb1*-GFP

The SLI selection system then requires selection with a different drug to select for parasites that have integrated the gene of interest into the genome. Thus, parasites that had taken up the plasmid episomally were treated with G418 consecutively for 14 days. No parasites were observed after 3 days of G418 drug pressure in either culture and were still not observed after 40 days of culturing (Figure 3-5). It can then be concluded that PfMyb1 is essential during the asexual stages of the parasite. This was expected as previous studies using the piggyBAC transposon have verified the essentiality of this gene in the asexual stages (67). Therefore, a conditional knock-down line had to be created to determine the essentiality in gametocyte stages.

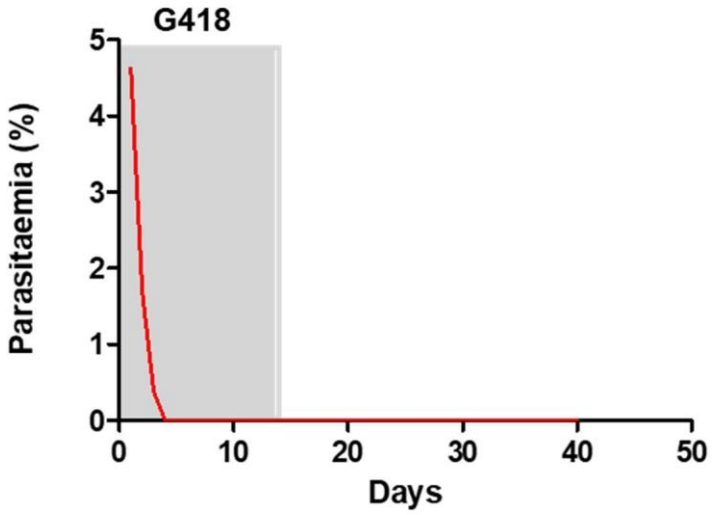

Figure 3-5. Growth of parasites during selection for integration. The grey bar indicates the length of treatment with G418 which was 14 days while the red line indicates parasitaemia. There were no observable parasites after 3 days of drug treatment. The culture was taken off drug pressure to allow the parasites to recover but there were still no observable parasites after 26 days of recovery.

3.1.6 Gene-specific primer design to amplify the 3' region of *pfmyb1*

Since the targeted gene disruption of *pfmyb1* did not yield recombinant parasites a conditional knock-down system was created. For this system, a 3' fragment of *pfmyb1* had to be cloned into the pSLI-*glms* (WT) and pSLI-*glms* (M9) plasmids. The MSF and MSR primers had a size of 36 and 33 bp (Table 3-2), respectively. The LiG homodimer for MSF is -15.4 kcal and MSR is -10.2 kcal. The LiG monomer for MSF is -0.2 kcal and -2.1 kcal for MSR and the LiG heterodimer is -6 kcal. The primers should produce a product size of 986 bp. Binding specificity was determined using Primer-BLAST (NCBI) and the primers were found to be specific and produce the *pfmyb1* product of 973 bp. There were two other products on a different gene that would produce fragments of >1700 bp and two products that would produce fragments of >3000 bp. Since the unwanted product sizes are very different from the intended product size, the primers could be used to amplify the 3' *pfmyb1* fragment.

Table 3-2. Properties of primers used to generate a 3' fragment of *pfmyb1.*

Name	Size (bp)	Tm (°C)	LiG homodimer @ 37 °C (kcal)	LiG monomer @ 37 °C (kcal)	LiG heterodimer @ 37 °C (kcal)	Product size (bp)	LiTm (°C)
MSF	36	62.1	-15.4	-0.2	-6	986	0.5
MSR	33	61.6	-10.2	-2.1			

3.1.7 Cloning *pfmyb1* into a conditional knock-down system

The amplification of the gene product (Figure 3-6A) yielded a band of approximately 1000 bp which corresponded to the theoretical size of 987 bp. The pSLI-*glms* (WT) plasmid and the control pSLI-*glms* (M9) plasmid were digested with the appropriate restriction enzymes. For

pSLI-*glms* (WT) (Figure 3-6B) two bands were produced, one of approximately 9 kbp and another of approximately 850 bp. The larger band corresponded to the backbone size (8138 bp) of pSLI-*glms* (WT). For pSLI-*glms* (M9), the same was produced and the larger band also corresponded to the backbone size of the plasmid (8138 bp).

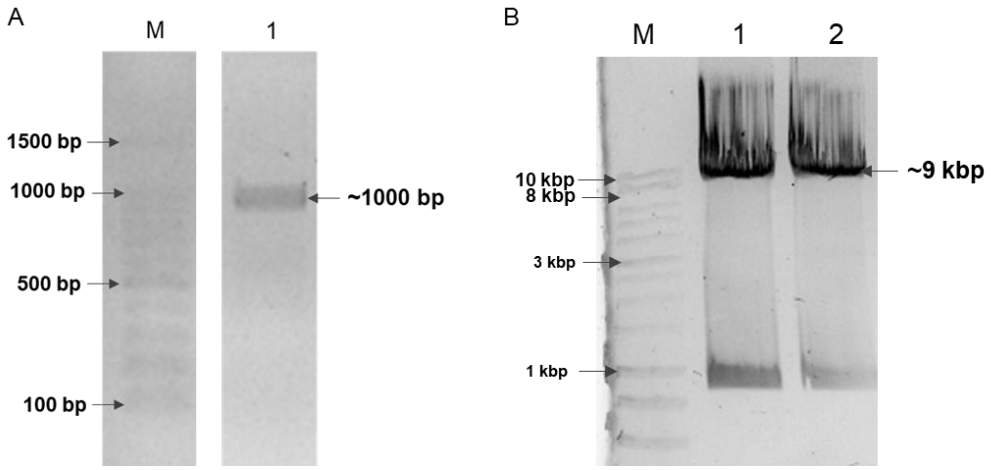

Figure 3-6. Cloning of a 3' *pfmyb1* gene fragment into pSLI-*glms* (WT) and pSLI-*glms* (M9). A. Amplification of the gene fragment with MSF and MSR primers. Lane M shows a 100 bp molecular marker and Lane 1 shows the PCR product of approximately 1000 bp. The reaction was visualised on a 1.5 % agarose/TAE gel poststained with EtBr **B.** Restriction enzyme digestion of pSLI-*glms* (WT) and pSLI-*glms* (M9) with *NotI* and *AvrII.* Lane M shows a 1 kbp marker. Lane 1 digestion of pSLI-*glms* (WT) yielded two bands of approximately 9 kbp and 850 bp that correspond to the backbone and insert size. Lane 2: digestion of pSLI-*glms* (M9), as with pSLI-*glms* (WT), shows two bands which corresponded to the backbone and insert size The reaction was visualised on a 1 % agarose/TAE gel post stained with EtBr.

The plasmid backbones were separately ligated with the amplified gene product, transformed and the colonies produced were screened for the presence of the pSLI-*pfmyb1*-gfp-glms and pSLI-*pfmyb1*-gfp-glms-mut plasmid by PCR. Gene-specific primers (MSF and MSR) were used for the PCR (Figure 3-7A). Positive colonies produced bands of approximately 1000 bp which corresponded to the expected size of 969 bp for both constructs. To confirm the correctness of the insert, the pSLI-*pfmyb1*-gfp-glms (Figure 3-7B) and pSLI-*pfmyb1*-gfp-glms-mut plasmids from the bacteria were sequenced (Figure 3-7C). The backbone, restriction sites and gene fragment were present without mutations when aligned with an *in silico* generated construct.

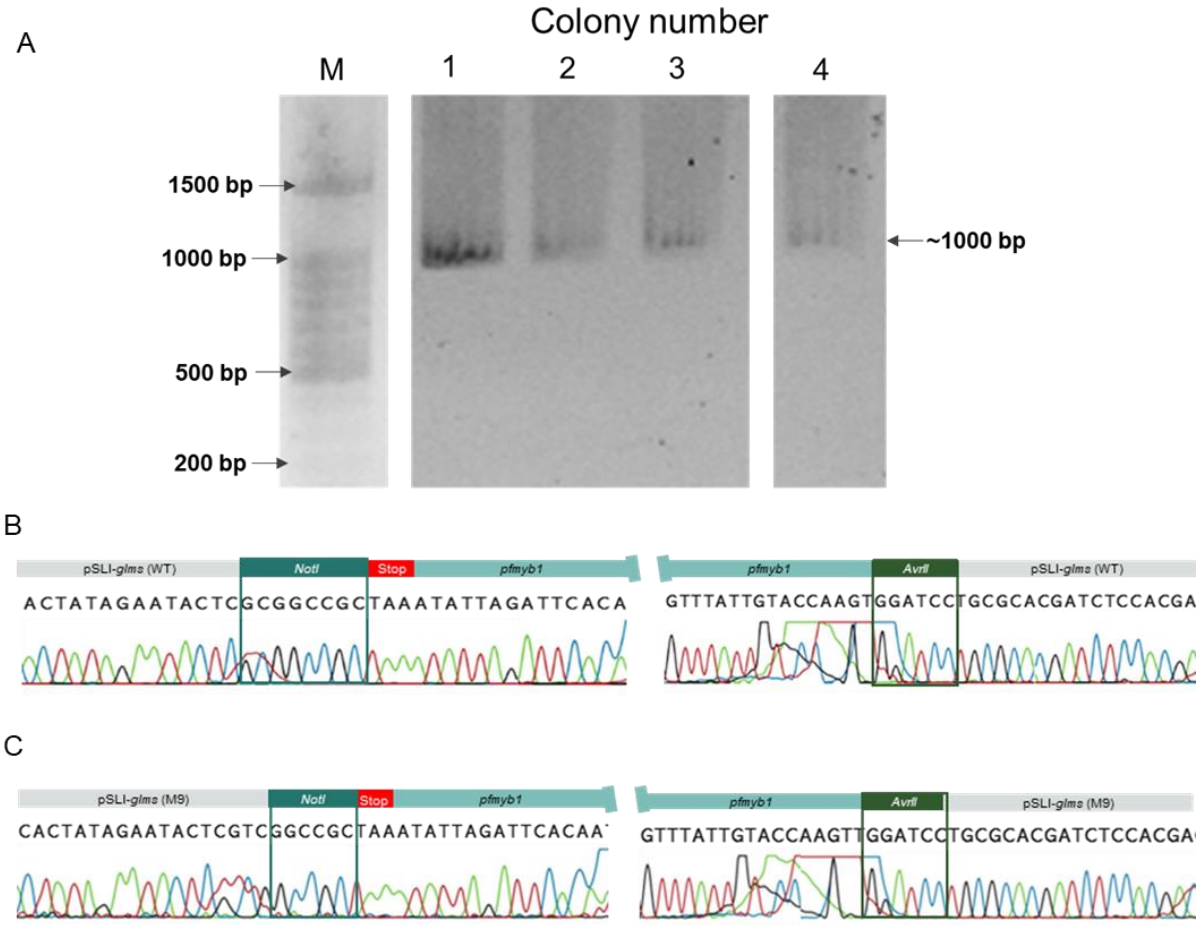

Figure 3-7. Confirmation of successful insertion of the 3' *pfmyb1* fragment into pSLI-*glms* (WT) and pSLI-*glms* (M9). A. A colony PCR was done on the resulting colonies after ligation. Lane M: 100 bp molecular marker. Lane 1-3: pSLI-*pfmyb*1-gfp-glms colonies which produced a band of approximately 1000 bp that corresponded to the theoretical size of 969 bp when using gene-specific primers. Lane 4: pSLI-*pfmyb1*-gfp-glms-mut had a single colony which produced a band of around 1000 bp that corresponded to the expected size of 969 bp. **B.** For pSLI-*pfmyb1*-gfp-glms colony 1 was sequenced and the resulting sequence showed the presence of the insert, backbone and both restriction sites without mutations. For the pSLI-*pfmyb1*-gfp-glms-mut colony was sequenced, and the resulting sequence showed the backbone, insert and both restriction sites without any mutations present.

3.1.8 Production of conditional knock-down lines for *pfmyb1*

The confirmed pSLI-*pfmyb1*-gfp-glms (68 µg) and pSLI-*Pfmyb1*-gfp-glms-mut (100 µg) plasmids were transfected into *P. falciparum* NF54 parasites using electroporation with time constants of 11.6 ms and 11.3 ms respectively. Even though the time constants were slightly below the lower limit of the optimal values, these should still be effective due to the large amount of plasmid used. The transfected parasites were treated with WR99210 for 10 days. Parasites from the pSLI-*pfmyb1*-gfp-glms transfected parasite culture were not observed after 3 days of treatment and were only observed after 33 days (Figure 3-8A). The pSLI-*pfmyb1*-gfp-glms-mut culture had no observable parasites after 4 days of drug treatment and were observed after 31 days (Figure 3-8A). The parasites were tested for the presence of episomal plasmid with the backbone specific primers SLI-F and GFP-R primers (Figure 3-8B). The parasite lines which recovered after selection, NF54-epi(pSLI-*pfmyb1*-gfp-*glms*) and

NF54-epi(pSLI-*pfmyb1*-gfp-glms-mut), both produced bands of approximately 1200 bp which corresponded to the expected size of 1178 bp which confirms the episomal uptake of the plasmid by the parasites.

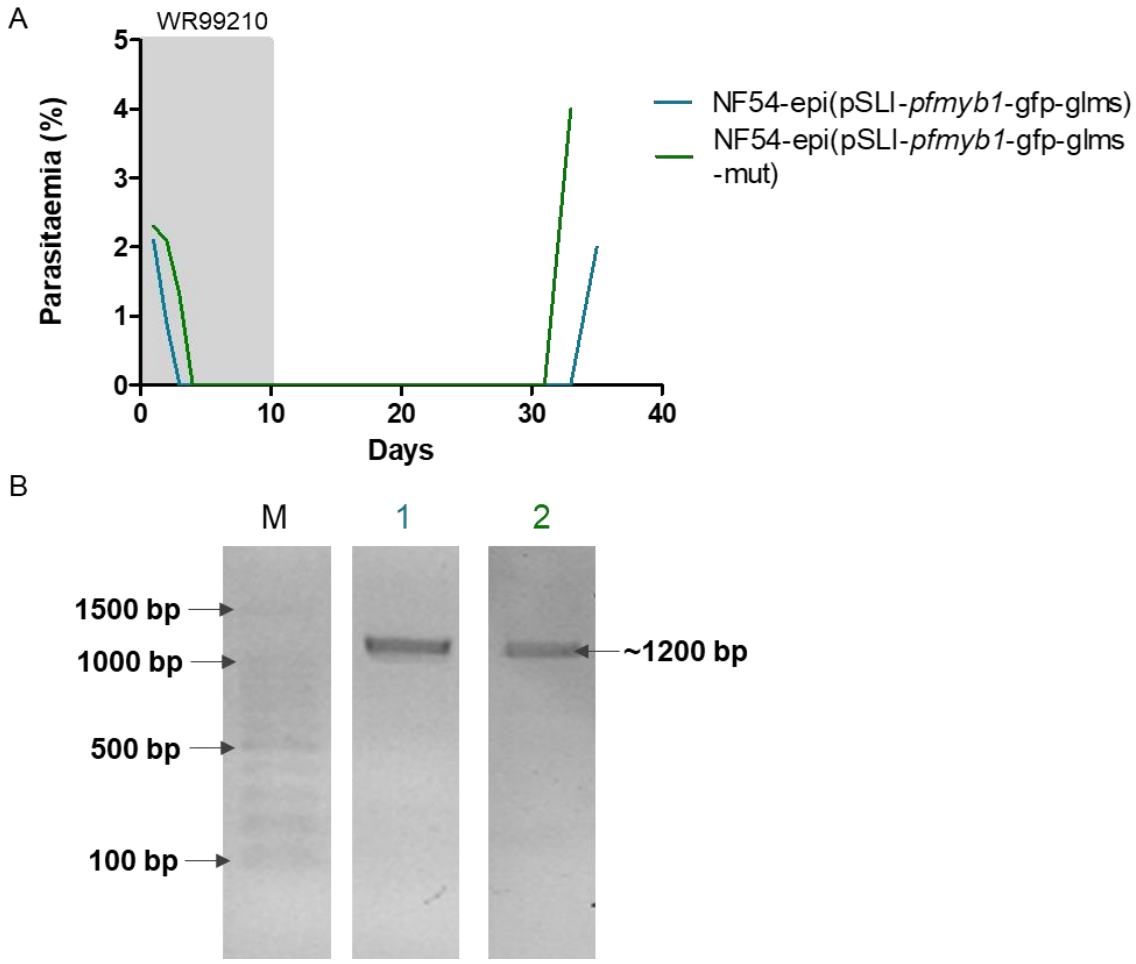

Figure 3-8. Production of episomal NF54-epi(pSLI-*pfmyb1*-gfp-*glms*) and NF54-epi(pSLI-*pfmyb1*-gfp-*glms-mut*) parasites. A. Parasites were treated with WR99210 for 10 days after transfection (as shown by the grey block). NF54-epi(pSLI-*pfmyb1*-gfp-*glms*) parasites were not observed after 3 days of drug pressure NF54-epi(pSLI-*pfmyb1*-gfp-glms-mut) parasites were not observed at 4 days. NF54-epi(pSLI-*pfmyb1*-gfp-*glms-mut*) parasites recovered first at 31 days and NF54-epi(pSLI-*pfmyb1*-gfp-*glms-mut*) parasites recovered after 33 days. **B.** Episomal uptake was determined by PCR with backbone specific primers. Lane M: 100 bp marker. Lane 1: NF54-epi(pSLI-*pfmyb1*-gfp-*glms*) produced a band of approximately 1200 bp. Lane 2: NF54-epi(pSLI-*pfmyb1*-gfp-*glms-mut)* produced a band of approximately 1200 bp which corresponded to the expected size of 1178 bp. The PCR product was visualised on a 1.5 % agarose/TAE gel poststained with EtBr.

Once episomal lines were confirmed, recombinant parasites were selected by treating with G418 for 14 days. NF54-epi(pSLI-*pfmyb1*-gfp-glms) parasites died after 4 days of treatment and NF54-epi(pSLI-*pfmyb1*-gfp-glms-mut) parasites died after 5 days of treatment. NF54-epi(pSLI-*pfmyb1*-gfp-glms-mut) parasites recovered after 22 days, and NF54-epi(pSLI-*pfmyb1*-gfp-glms) parasites recovered after 25 days (Figure 3-9A). To determine integration of the recombinant *pfmyb1*-fragment into the genome, the parasites were tested on three criteria: presence of the wildtype *pfmyb1* locus, integration of the *pfmyb1* gene fragment with the glms ribozyme sequence and GFP in the 5' region of the gene and integration of the full plasmid length with the original gene on the 3' end (Figure 3-9B). If there is recombination,

then the distance between the primers is too large and no product will be produced, therefore, a band will be produced if no integration occurred. The wildtype *pfmyb1* locus only showed a band in WT NF54 parasites of approximately 1750 bp (Figure 3-9C) which corresponded to thetheoretical size of 1758 bp (Figure 3-9D). For 5' integration, no product will be produced if no recombination occurred as there will not be a GFP sequence present. A band of approximately1600 bp showed in both NF54-epi(pSLI-*pfmyb1*-gfp-glms) and NF54-epi(pSLI-*pfmyb1*-gfp- glms-mut*)* when testing for 5' integration (Figure 3-9C) which corresponded to the expected size of 1654 bp (Figure 3-9D). For 3' integration bands of approximately 1200 bp was produced for both NF54-epi(pSLI-*Pfmyb1*-gfp-*glms-mut*) and NF54-epi(pSLI-*Pfmyb1*-gfp- *glms*) which corresponded to the expected size of 1259 bp. Again, since the SLI-F sequence isplasmid backbone specific there will be no product produced if there is no integration. Since this is a wildtype parasite line, no recombination was expected. Thus, it was used as a negative control.

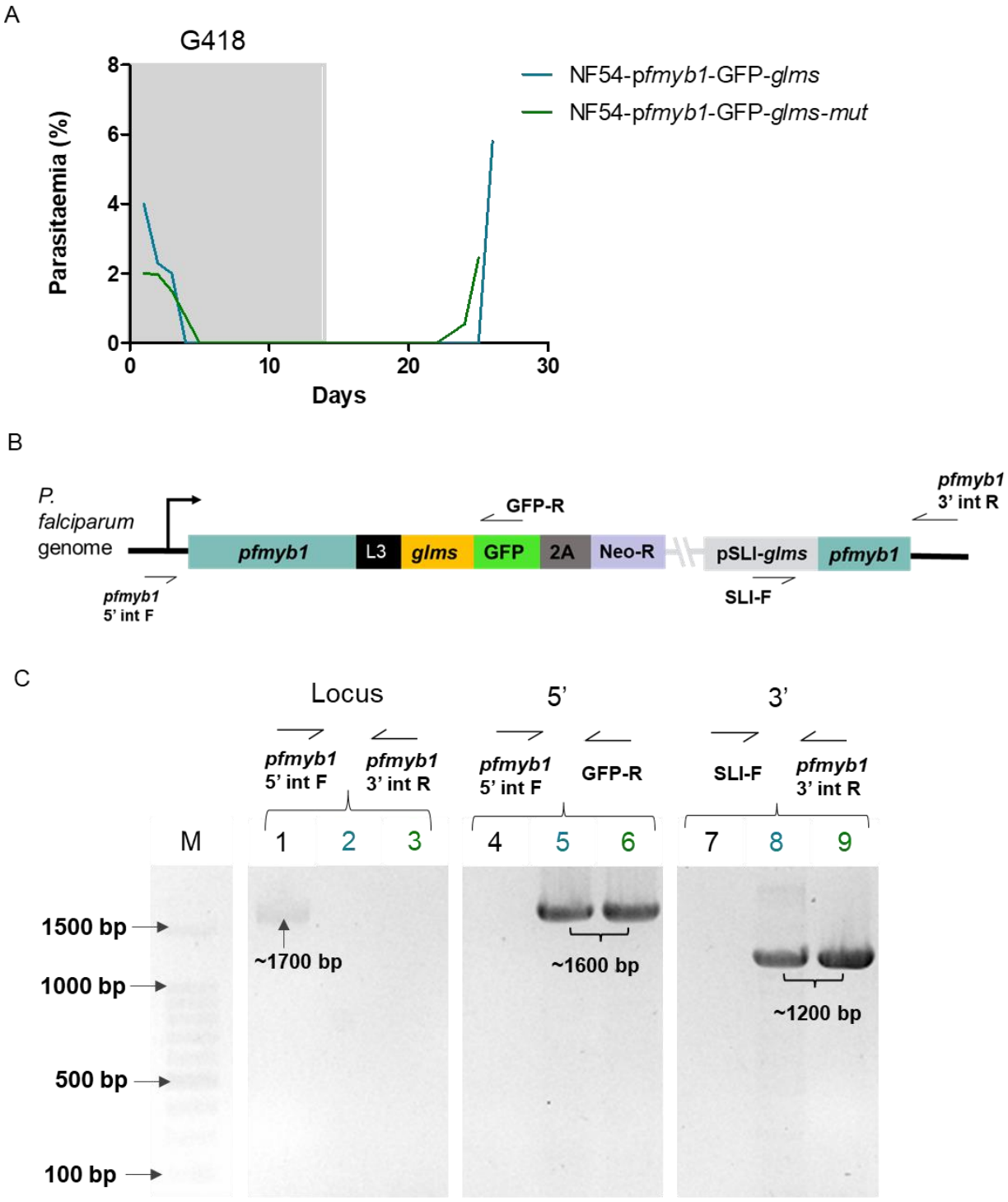

Figure 3-9. Production of the conditional knock-down line NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut*. A. Time taken for the parasitaemia to be unobservable (3 days after drug pressure was started) and recover (29 days after drug pressure was started, 15 days after drug pressure was stopped). **B.** A schematic showing the approximate locations of the primers used. A5' int F and ATGD 3' R was used to test for whether there was integration into the locus. The 5' PCR was used to test the correct integration of the gene fragment that is connected to the GFP tag using the A 5' int F and GFP-R primer. The 3' integration PCR was used to test whether the remnants of the gene were present next to the backbone of the plasmid. This was done using the SLI-F and ATGD-3' R primers. **C.** A PCR showing whether the parasites were integrated and free from WT. Lane M: A 100 bp molecular marker. Lane 1: Locus PCR for NF54. This was used as a positive control and as expected, a band of approximately 1700 bp was observed. Lane 2: Locus PCR for NF54-*pfmyb1*-GFP-*glms*. No bands were observed. Lane 3: Locus PCR for NF54-*pfmyb1*-GFP-*glms-mut*. No band was observed. Lane 4: 5' integration PCR for NF54. No band was observed. Lane 5: 5' integration PCR for NF54-*pfmyb1*-GFP-*glms*. A band of approximately 1600 b was observed which corresponded to the expected size of 1654 bp. Lane 6. 5' integration for NF54-*pfmyb1*-GFP-*glms-mut*. A band of approximately 700 bp was observed which corresponded to the expected size of 658 bp.

Lane 7. 3' integration PCR for NF54. No band was observed. Lane 8: 3' integration PCR for NF54-*pfmyb1*-GFP-*glms*. A band of 539 bp was expected but no band was observed. Lane 9: 3' integration PCR for NF54-*pfmyb1*-GFP-*glms-mut*. A band of 539 bp was expected, but no band was observed. The reactions were visualised on a 1.5 % agarose/TAE gel stained with EtBr.

3.1.9 Validation of conditional knock-down line

Once the recombinant *pfmyb1* gene fragment was present in the culture without any wildtype present the culture had to be verified as expressing the recombinant gene. This was done by using the GFP tag through blotting techniques. A dot blot was done with separated nuclear and cytoplasmic fractions to try and reduce non-specific binding by the antibody (Figure 3-10A). The NF54-*pfmyb1*-GFP-*glms* and NF54-*Pfmyb1*-GFP-*glms-mut* lines did not show an increase in fluorescence compared to the NF54 line as shown by the integrated density, when normalised against HSP70 (Figure 3-10B). This may be due to several possibilities. The recombinant protein of interest had degraded during isolation, there was non-specific binding of the GFP antibody, and the recombinant protein is not highly expressed. To combat these problems, it was decided to do a western blot.

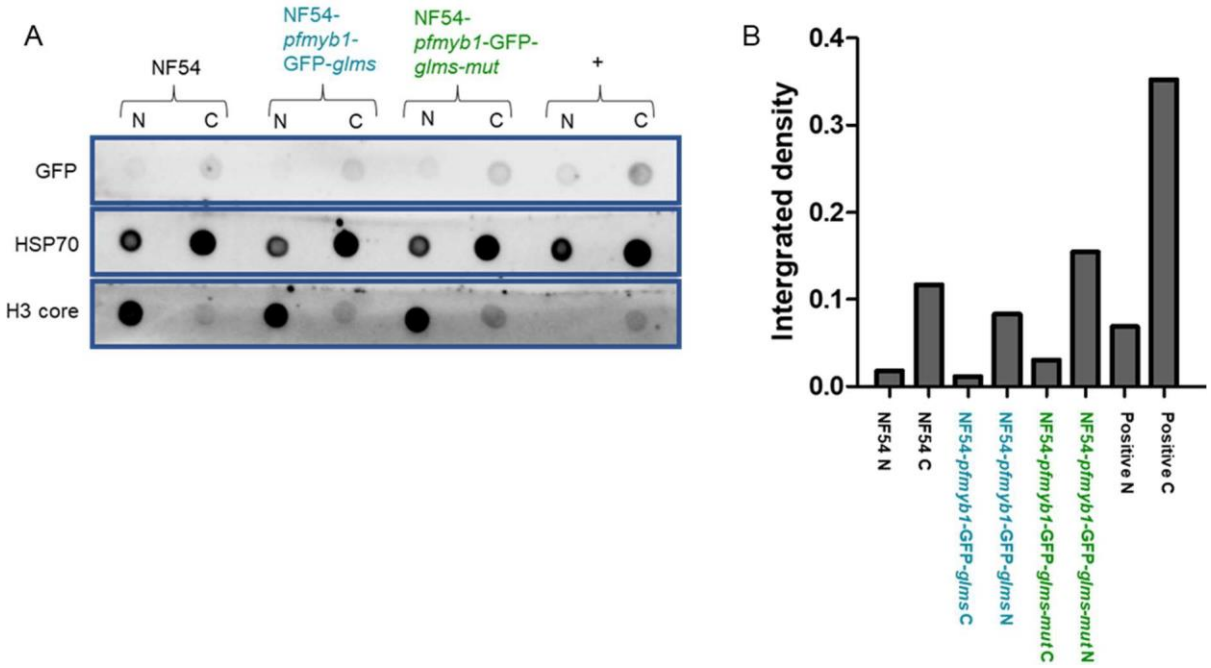

Figure 3-10. Validation of the NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines for GFP expression. A. A dot blot with both nuclear (N) and cytoplasmic (C) fractions. The dots were tagged with anti-GFP antibodies, anti-HSP70 antibodies and anti-H3 core antibodies. GFP did not produce a strong signal except for the cytoplasmic fraction of the positive control. As expected Hsp70 showed a higher signal in the cytoplasmic fraction as compared to the nuclear fraction. H3 core showed a higher signal intensity in the nuclear fractions. **B.** An integrated density chart of the normalised density of the anti-GFP tagged dots. As can be seen, the fluorescence produced by the NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines was not as high as the positive control but was more comparable to the NF54.

The western blot showed a band of 78 kDa in both the NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines which corresponded to the expected size, but it showed a band in NF54 as well (Figure 3-11A). The loading control, H3 core, was not visible for all three samples

(Figure 3-11B) which suggests an uneven amount of protein was loaded between samples. This would affect the intensity of the signal reported. Thus, it was unknown whether recombinant GFP was being visualised and since a band was visible in NF54 which was the same size as the protein of interest, it was determined that there was non-specific binding of the GFP-antibody, but this data remains inconclusive.

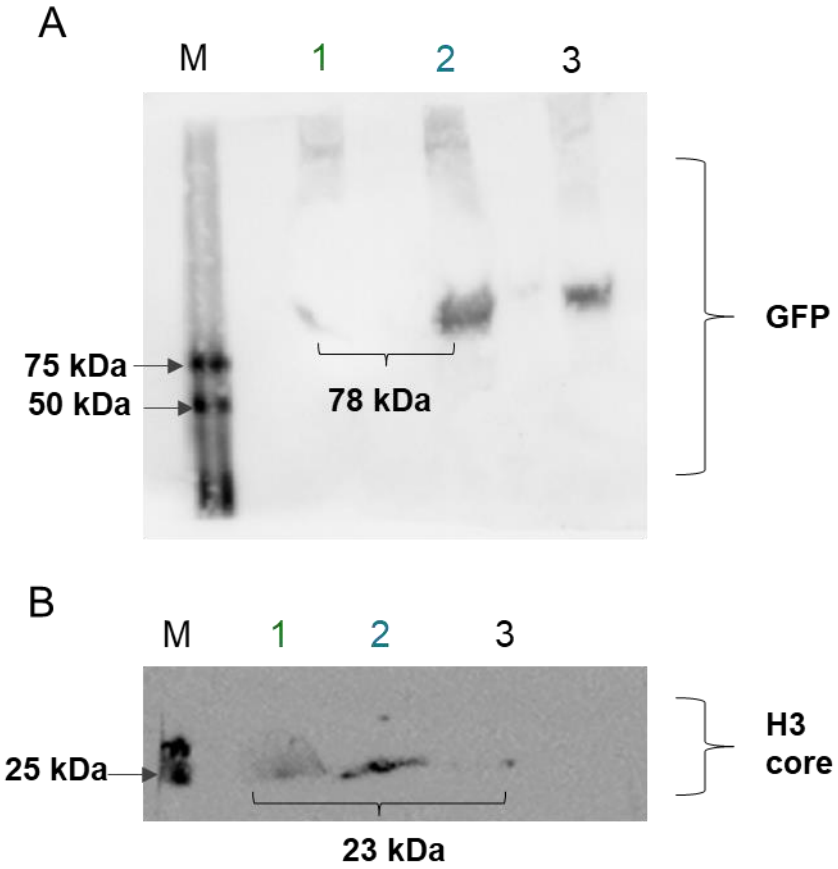

Figure 3-11. Western blot of NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines to determine GFP expression. A. The membrane was incubated with GFP antibodies. Lane M: Protein marker. Lane 1: NF54-*Pfmyb1*-GFP-*glms-mut* faintly shows a band of approximately 78 kDa even though the image is not clear due to an air bubble which was present during the transfer. Lane 2: NF54-*pfmyb1*-GFP-*glms* showed a band of approximately 78 kDa. Lane 3: NF54 incubated with anti-GFP antibodies showed a band of approximately 78 kDa. **B.** Loading control using H3 core antibodies. Lane M: Protein marker, Lane 1: NF54-*pfmyb1*-GFP-*glms-mut* shows a band of approximately 23 kDa that corresponded to the expected size of H3 core. Lane 2: NF54-*pfmyb1*-GFP-*glms* showed a band of approximately 23 kDa. Lane 3: NF54 shows an extremely faint band at approximately the size of the other bands.

Alternatively, GFP expression by the recombinant parasites was observed using confocal microscopy (Figure 3-12). The parasites were co-stained with Hoechst to stain the parasite nucleus and verify the presence of the parasite. There is a clear GFP signal as well as a Hoechst signal signifying that the parasites are expressing GFP and are likely recombinant since they are expressing the reporter gene.

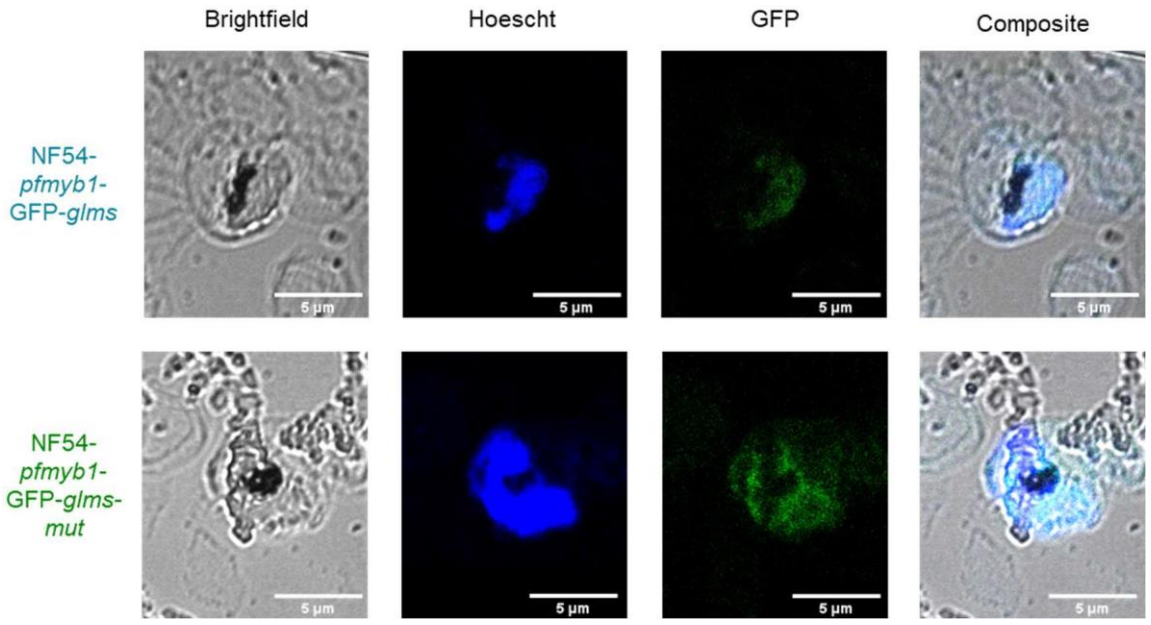

Figure 3-12. Fluorescence microscopy of NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines to visualise GFP expression. A majority trophozoite culture was used as this is the stage where PfMyb1 is most highly expressed. Hoechst was positive for the parasite nucleus and GFP was expressed and localised to the Hoechst in the nucleus.

Once the lines were verified as expressing GFP, they were compared to the NF54 WT to ensure the gene recombination did not have any unwanted effects. The proliferation between the triplicate repeats of the two recombinant lines and a WT NF54 line was observed. During the asexual stages NF54-*pfmyb1*-GFP-*glms* ($p=0.7483$, $p<0.05$) and NF54-*pfmyb1*-GFP-*glms-mut* ($p=0.9824$, $p<0.05$) did not show any statistical difference in parasitaemia over 5 days when compared to a WT NF54 line (Figure 3-13A). There were also no observable differences in morphology between the NF54 strain and the two recombinant lines (Figure 3-13B).

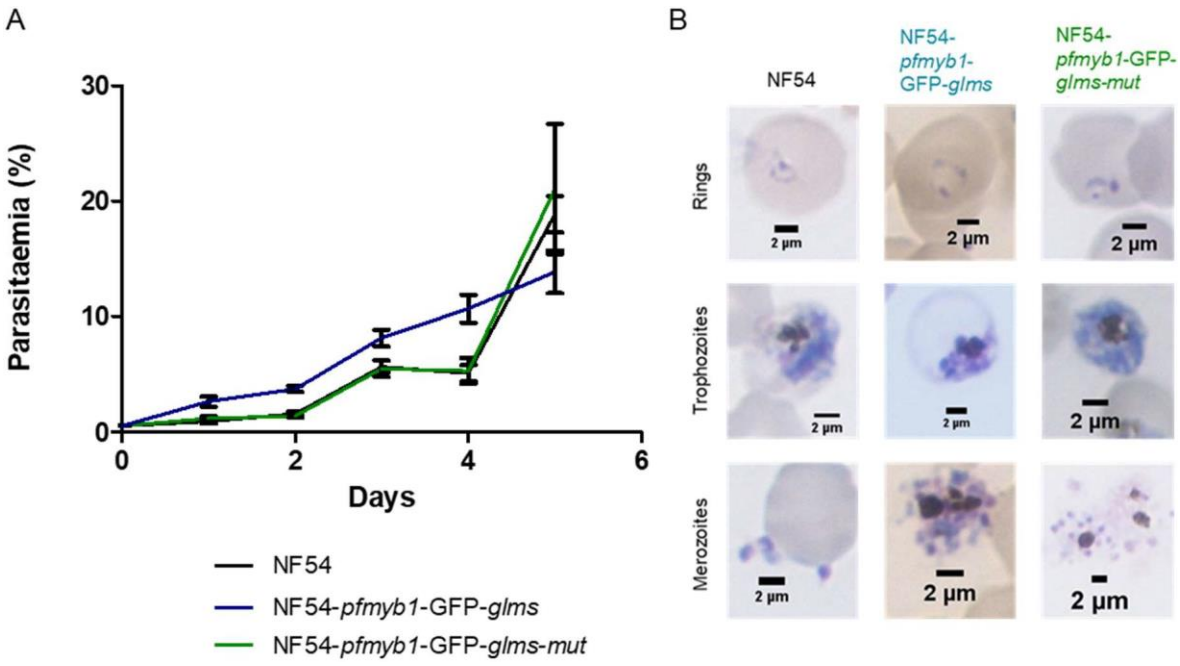

Figure 3-13. Validation of the NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines' asexual stages using proliferation and morphology. A. Asexual growth analysis of NF54-*Pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* when compared to NF54. NF54-*Pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* proliferation did not show any significant difference when compared to NF54. **B.** Morphological analyses of recombinant parasites compared to WT parasites. As can be seen there are no morphological differences of the recombinant lines as compared to the wild-type NF54 line.

The gametocyte stages were also observed. During the asexual stages, before gametocytogenesis, NF54-*pfmyb1*-GFP-*glms* ($p=0.5978$, $p<0.05$) and NF54-*pfmyb1*-GFP-*glms-mut* ($p=0.7164$, $p<0.05$) showed no statistical difference when compared to WT NF54 parasites (Figure 3-13A, dotted lines). During the gametocyte stages NF54-*pfmyb1*-GFP-*glms* ($p=0.7418$, $p<0.05$) and NF54-*pfmyb1*-GFP-*glms-mut* ($p=0.9164$, $p<0.05$) did not show any statistical difference in parasitaemia and gametocytaemia over 13 days (Figure 3-14A, solid lines) when compared to a WT NF54 line. There were also no observable differences in morphology between the lines (Figure 3-14B).

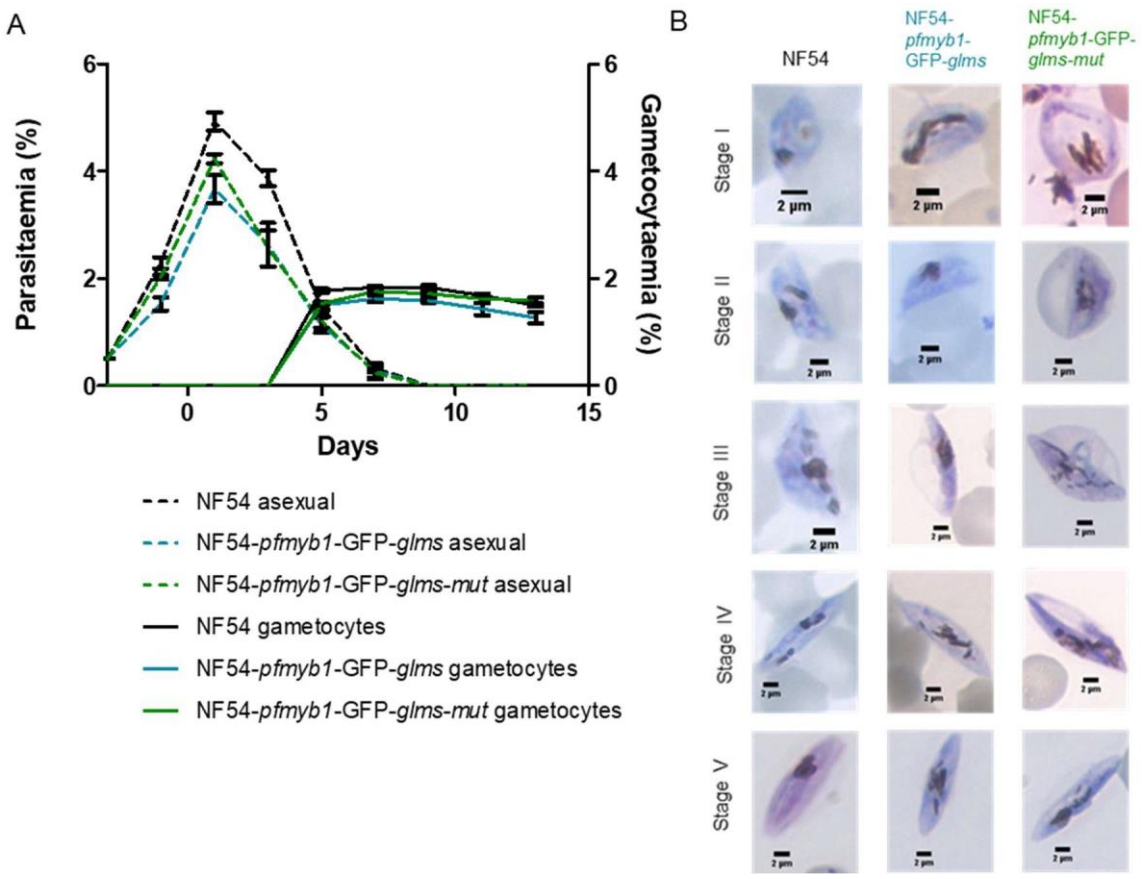

Figure 3-14. **Gametocyte growth and morphology analysis of NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* when compared to NF54**. **A.** The dotted lines show the parasitaemia and the solid lines show the gametocytaemia over 13 days. **B.** Images of the sexual stages of the recombinant parasites. As can be seen here, there are no distinguishable morphological differences between the recombinant lines and the WT NF54 line.

Thus, PfMyb1 is essential for asexual stages as determined by TGD. Conditional knock-down lines were generated which were found to proliferate and differentiate comparably to WT NF54 parasites.

3.2 Pf3D7_0603600

3.2.1 *In silico* domain classification of Pf3D7_0603600

According to Interpro, Pf3D7_0603600 was found to be part of the ARID-DBD superfamily. This family consists of proteins with an ARID DBD and thus Pf3D7_0603600 can be thought to be a putative TF. The ARID DBD is approximately 100 residues and consists of six α-helices separated by turns. A homology model of the ARID DBD was obtained by using the SWISS-model repository.

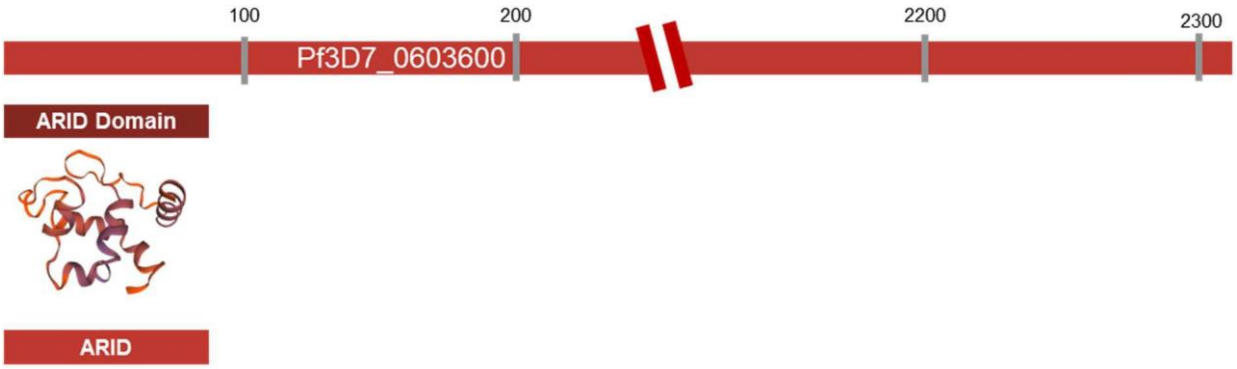

Figure 3-15. *In silico* analysis of PF3D7_0603600. PF3D7_0603600 falls under the ARID-DBD superfamily and contains an ARID DBD from residue 1 to 95.

3.2.2 Gene-specific primer design to amplify the 5' region of *pf3d7_0603600*

The ATF and ATR primers had a size of 35 and 33 bp (Table 3-3) respectively and were designed to disrupt the ARID DBD of *pf3d7_0603600.* Since the functional domain was smaller than the recommended size for TGD, the ATF primer started at +57 bp and due to the stop codon inserted by the primer, gene disruption will occur. The LiG homodimer for ATF was -15.4kcal and ATR was -7.7 kcal. The LiG monomer for ATF was 0 kcal and -0.5 kcal for ATR and the LiG heterodimer was -4.9 kcal. The primers should produce a product size of 417 bp.

Table 3-3. Properties of primers used to amplify the 5' fragment of *pf3d7_0603600*

Name	Size (bp)	Tm (°C)	LiG homodimer @ 37 °C (kcal)	LiG monomer @ 37 °C (kcal)	LiG heterodimer @ 37 °C (kcal)	Product size (bp)	LiTm (°C)
ATF	35	63.1	-15.4	0	-4.9	417	2.0
ATR	33	61.1	-7.7	-0.5			

3.2.3 Cloning the 5' *pf3d7_0603600* gene fragment into the pSLI-TGD system.

To test essentiality of Pf3D7_0603600 in asexual parasites, a 5'gene fragment was cloned into a knockout system. First a 5' gene fragment had to be produced. This was done using the ATF forward primer and the ATR reverse primer (Figure 3-16A). A band of approximately 450bp was produced which corresponded to the expected size of 417 bp. The pSLI-TGD plasmid was digested with *NotI* and *MluI* to prepare the backbone for directional cloning. Two bands of approximately 7 kbp and 2 kbp were produced. The larger of the bands corresponded to theexpected size of the pSLI-TGD backbone of 6783 bp.

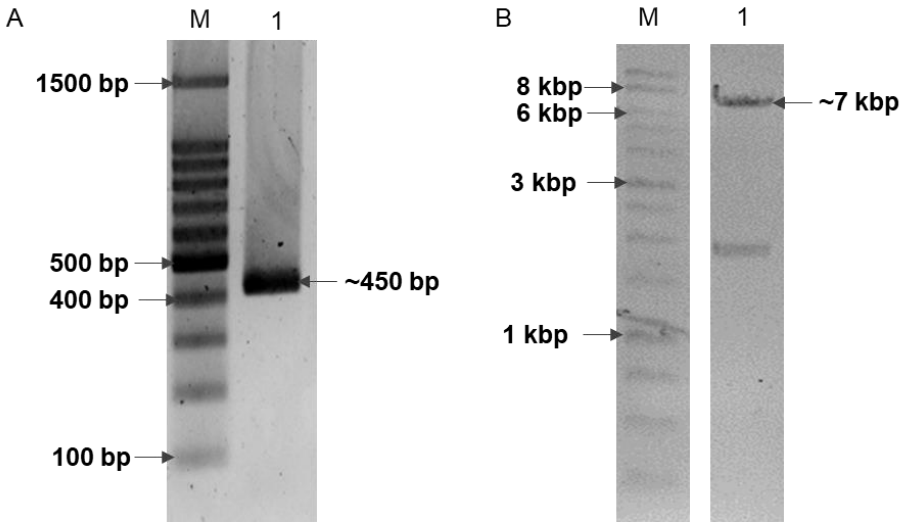

Figure 3-16. Cloning of a 5' *pf3d7_0603600* fragment into the pSLI-TGD vector. A. The *Pf3D7_0603600* 5'gene fragment was produced using the ATF and ATR gene-specific primers. Lane M: 0.5 µg of a 100 bp DNA ladder. Lane 1 shows a band of approximately 450 bp was produced which corresponded to the expected size of 417 bp. The PCR was visualised on a 1.5 % agarose/TAE gel post-stained with EtBr. **B.** Restriction enzyme digestion of pSLI-TGD with *NotI and MluI* to produce sticky ends for directional cloning. Lane M: 0.5 µg of a 1 kbp ladder. Lane 1 shows two bands of ~ 7 kbp and ~ 1.8 kbp.

Since both PCRs yielded the expected band sizes the 5' gene fragment of *pf3d7_0603600* was ligated into pSLI-TGD and transformed into bacteria. The resulting colonies were screened forthe presence of the plasmid by colony PCR using the gene-specific primers ATF and ATR. All six colonies screened showed a single band of approximately 450 bp (Figure 3-17A) which corresponded to the expected size of 417 bp. Colony 5 was sequenced (Figure 3-17B), the *NotI* and *MluI* restriction sites were present, and no mutations were observed when the sequencing data was aligned to an *in silico* generated construct.

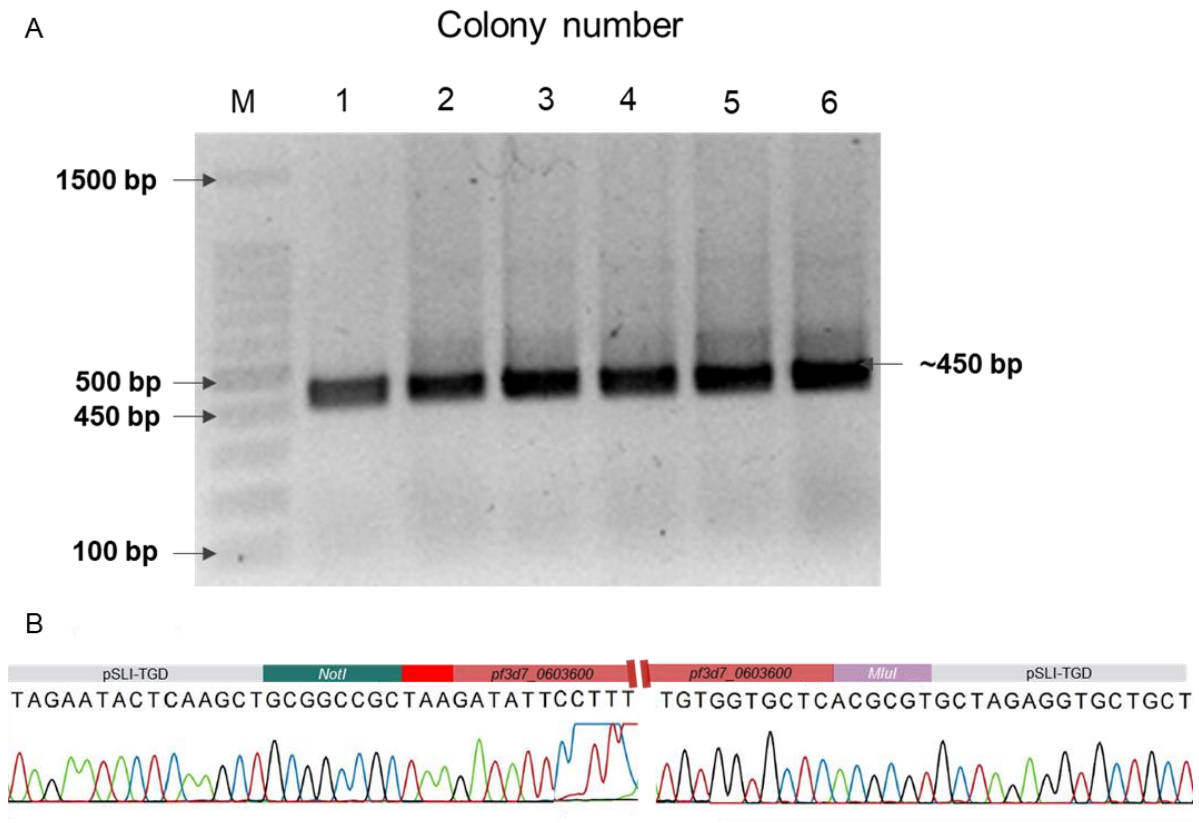

Figure 3-17. Production of the pSLI-*pf3d7_0603600*-gfp plasmid. A. Colonies were screened by using gene-specific primers. Lane M: A 100 bp ladder (0.5 µg). Lane 1-6: Colonies 1-6. All 6 colonies that were tested showed a band of approximately 450 bp. This corresponded to the expected size of 417 bp. **B.** Colony 5 was sequenced and the plasmid backbone, both restriction sites and the insert were observed.

3.2.4 Production of NF54-epi(pSLI-*pf3d7_0603600*-gfp)

Once the pSLI-*pf3d7_0603600*-gfp plasmid was verified, it was prepared for transfection. A total of 100 µg of plasmid was transfected into a majority ring stage NF54 culture with a time constant of 11.7 ms. The transfected parasites were drugged with WR99210, and the parasites were not observed using microscopy after 3 days. They were drugged for another 10 days resulting in 13 days on WR99210. The parasites recovered after 29 days (19 days after the drug was removed) (Figure 3-18A). DNA was isolated to determine the presence of the plasmid episomally. This was done using a PCR with plasmid backbone specific primers (SLI-F and GFP-R) and a band of approximately 500 bp was observed. This corresponded to the expected size of 578 bp (Figure 3-18B).

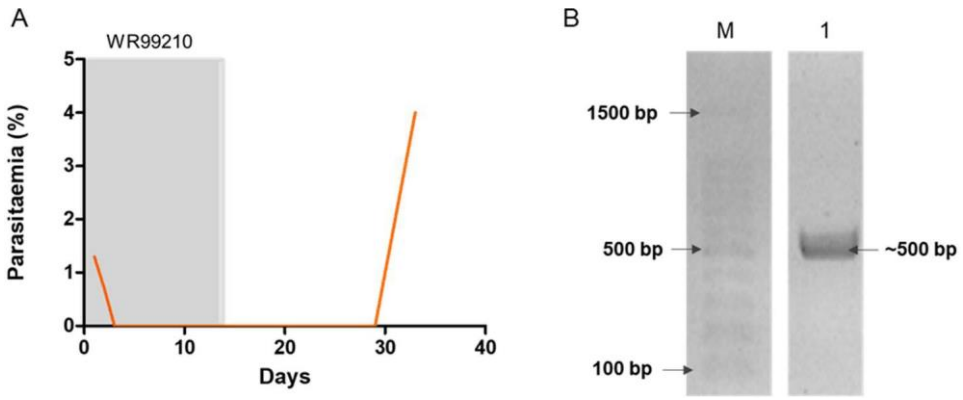

Figure 3-18. Production of NF54-epi(pSLI-*pf3d7_0603600*-gfp) lines through episomal uptake of plasmids. **A.** The time taken from transfection to recovery of episomal parasites. **B.** To confirm episomal uptake of the plasmid into the parasites a PCR was done using plasmid backbone specific primers. Lane M: a 100 bp molecular marker. Lane 1: gDNA isolated from the transfected parasites used in a PCR showed a band of approximately 500 bp. The reaction was visualised on a 1.5 % agarose/TAE gel poststained with EtBr.

Once episomal uptake was confirmed, the parasites were treated with G418 for 14 days to select for parasites that had integrated the recombinant gene fragment into the genome. The parasites were not observed after 3 days in all three cultures and replicate one and two were observed after 14 days of drug pressure and 15 days recovery (Figure 3-19A). The parasitaemia of the two replicates were averaged for each day. A PCR was done to determine integration on the recovered parasites. The locus PCR is used to determine whether the gene in the parasite's genome recombined with the gene fragment in the plasmid. Since the integration primers span the gene fragment if recombination occurred the primers would be too far apart for the polymerase to replicate the DNA effectively. Thus, no bands are expected in the cultures which have undergone drug selection and recovered (Figure 3-19B). Bands were present in all three locus PCRs (Figure 3-19C) meaning that WT parasites were still present in the treated cultures. The 5' integration PCR is used to determine whether the gene fragment of interest has recombined with the parasite genome and the GFP tag is still present and in the correct place. WT NF54 parasites that do not contain the pSLI-*pf3d7_0603600*-gfp plasmid was used as a negative control. No band was expected because the GFP primer sequence is present on the pSLI-TGD backbone (Figure 3-19B). No band was observed for the NF54 WT parasites as expected. Products produced by the 5' integration PCR band was present in both replicates indicating successful recombination. However, no bands for the 3' integration PCR were observed even though a band of 539 bp was expected for replicate one and two. This indicates that the parasites are not recombinant which conflicts with the results of the 5' PCR.

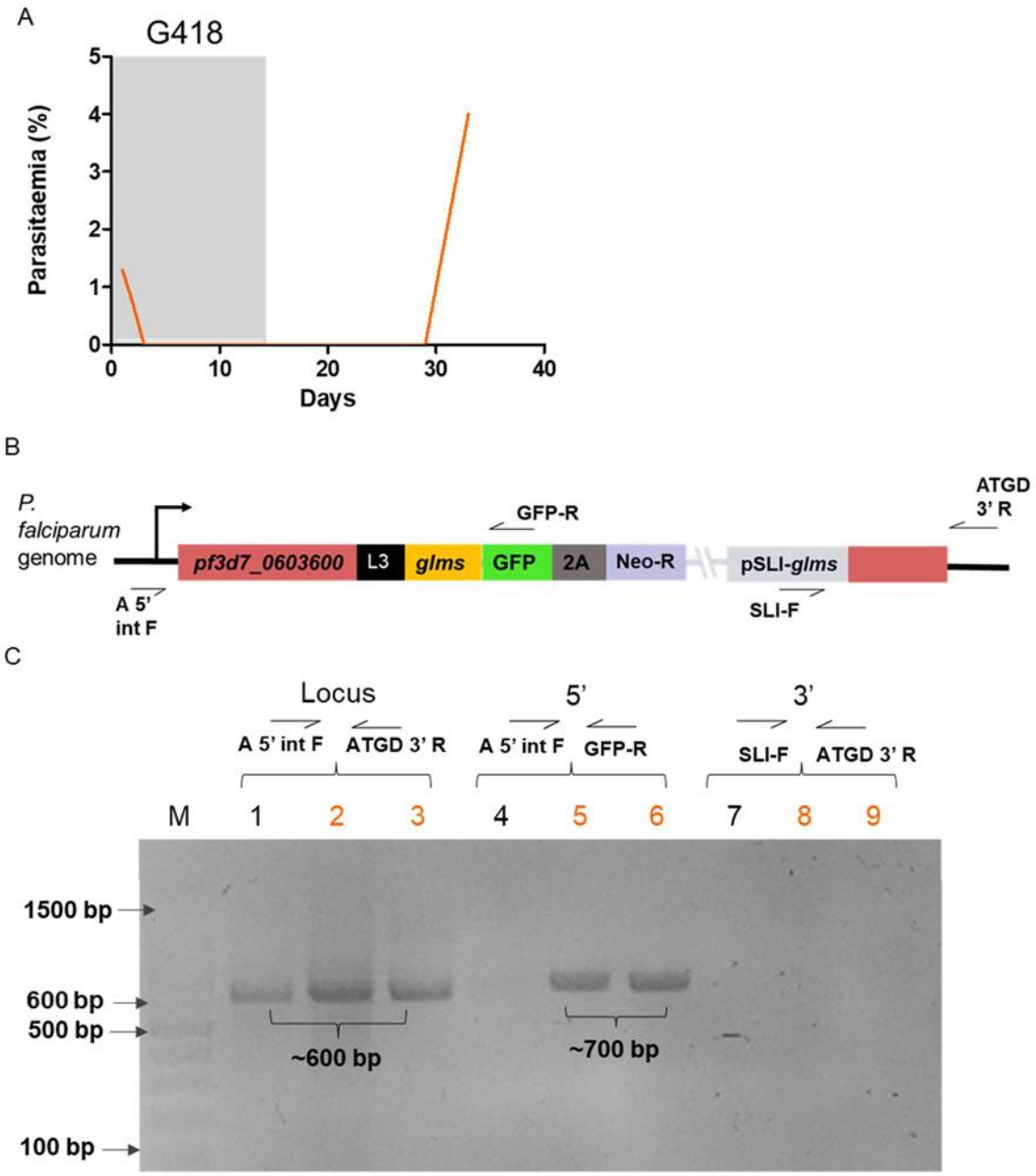

Figure 3-19. Selection for NF54-*LIpf3d7_0603600*-GFP integrated parasites. A. Time taken for the parasitaemia to be unobservable (3 days after drug pressure was started) and recover (29 days after drug pressure was started, 15 days after drug pressure was stopped). **B.** A schematic showing the approximate locations of the primers used. A5' int F and ATGD 3' R was used to test for whether there was integration into the locus. If there is recombination, then the distance between the primers is too big and no product will be produced. If there is no integration there will be a band. The 5' PCR was used to test the correct integration of the gene fragment that is connected to the GFP tag using the A 5' int F and GFP-R primer. The 3' integration PCR was used to test whether the remnants of the gene were present next to the backbone of the plasmid using the SLI-F and ATGD-3' R primers. **C.** A PCR showing whether the parasites were integrated and free from WT. Lane M: A 100 bp molecular marker. Lane 1: Locus PCR for NF54. This was used as a positive control and as expected, a band of approximately 600 bp was observed. Lane 2: Locus PCR for replicate one. No band was expected here but a 600 bp band was observed. Lane 3: Locus PCR for replicate two. No band was expected but a band of approximately 600 bp was observed. Lane 4: 5' integration PCR for NF54. Since this is a WT parasite line, no recombination was expected. Thus, it was used as a negative control. No band was expected, and none was observed. Lane 5: 5' integration PCR for replicate one. A band of approximately 700 b was observed which corresponded to the expected size of 658 bp. Lane 6. 5' integration for replicate two. A band of approximately 700 b was observed which corresponded to the expected sizeof 658 bp. Lane 7. 3' integration PCR for NF54. There is no integration present, thus no band was expected. No band was observed. Lane 8: 3' integration PCR for replicate one. A band of 539 bp was expected but no band was

observed. Lane 9: 3' integration PCR for replicate two. A band of 539 bp was expected, but no band was observed. The reactions were visualised on a 1.5 % agarose/TAE gel stained with EtBr.

3.2.5 Gene-specific primer design to amplify the 3' region of *pf3d7_0603600*

A conditional knock-down system was produced for *pf3d7_0603600.* A 3' fragment of *pf3d7_0603600* had to be cloned into the pSLI-*glms* (WT) and pSLI-*glms* (M9). The primers designed to amplify this region are ASF (33 bp) and ASR (39 bp). The properties are summarised in Table 3-4.

Table 3-4. Properties of the two primers used to produce a 3' fragment of *pf3d7_0603600.*

Name	Size (bp)	Tm (°C)	LiG homodimer @ 37 °C (kcal)	LiG monomer @ 37 °C (kcal)	LiG heterodimer @ 37 °C (kcal)	Product size (bp)	LiTm (°C)
ASF	33	63.1	-15.4	-0.2	-8.3	979	2.0
ASR	39	61.1	-7.7	-2.1			

3.2.6 Cloning for conditional knock-down systems

The 3' end of the gene was amplified (Figure 3-20A) and a band of approximately 950 bp was observed which corresponded to the expected band size of 969 bp. pSLI-*glms* (WT) and pSLI-*glms* (M9) were digested with *NotI* and *MluI* to create sticky ends for cloning. A band of approximately 8 kbp was observed which corresponded to the expected size (Figure 3-20B)

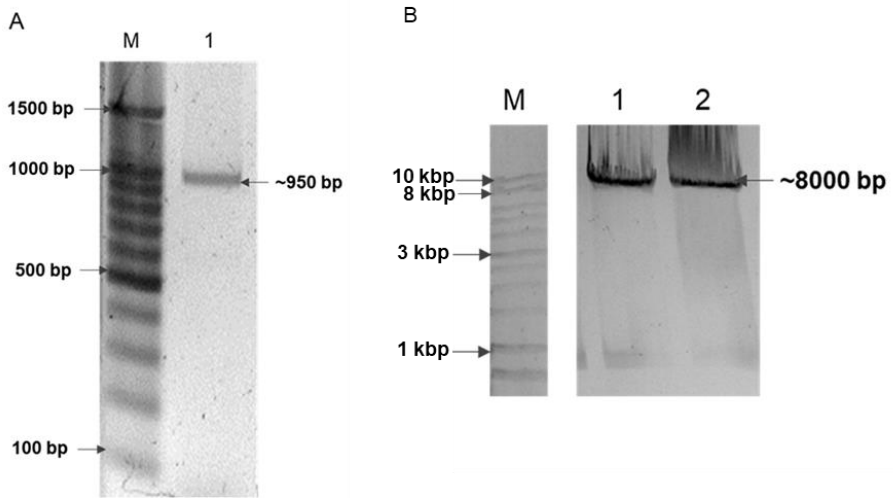

Figure 3-20. Production of a conditional knock-down system for PF3D7_0603600. A. A 3' gene fragment was amplified using gene-specific primers shown in A. Lane M: 0.5 µg of a 100 bp molecular ladder. Lane 1: A band of approximately 950 bp was observed which corresponded to the expected size of 969 bp. The reaction was visualised on a 1.5 % agarose/TAE gel post stained with EtBr. **B.** The pSLI-*glms* (WT) and pSLI-*glms* (M9) plasmids were digested with the appropriate restriction enzymes to create sticky ends for directional cloning. Lane M: a 1 kbp marker. Lane 1: pSLI-*glms* (WT) was digested with *NotI* and *AvrII* and a band of approximately 8000 bp was observed. Lane 2: pSLI-*glms* (M9) was digested with *NotI* and *AvrII* and a band of approximately 8000 bp was observed. The reaction was visualised on a 1 % agarose/TAE gel poststained with EtBr.

The 3' gene fragment and plasmid backbone were ligated and transformed into bacteria. The resulting colonies were screened using colony PCR with gene-specific primers. Of the colonies screened, two from pSLI-*pf3d7_0603600*-gfp-glms and three from pSLI-*pf3d7_0603600*-gfp-glms-mut showed a band of approximately 1000 bp which corresponded to the expected size of 969 bp (Figure 3-21A). One colony from both pSLI-*pf3d7_0603600*-gfp-glms (Figure 3-21B) and from pSLI-*pf3d7_0603600*-gfp-glms-mut (Figure 3-21C) was sequenced. For both constructs the backbone, insert and both restriction sites are observed and alignment to the *in silico* generated construct showed that no mutations were present.

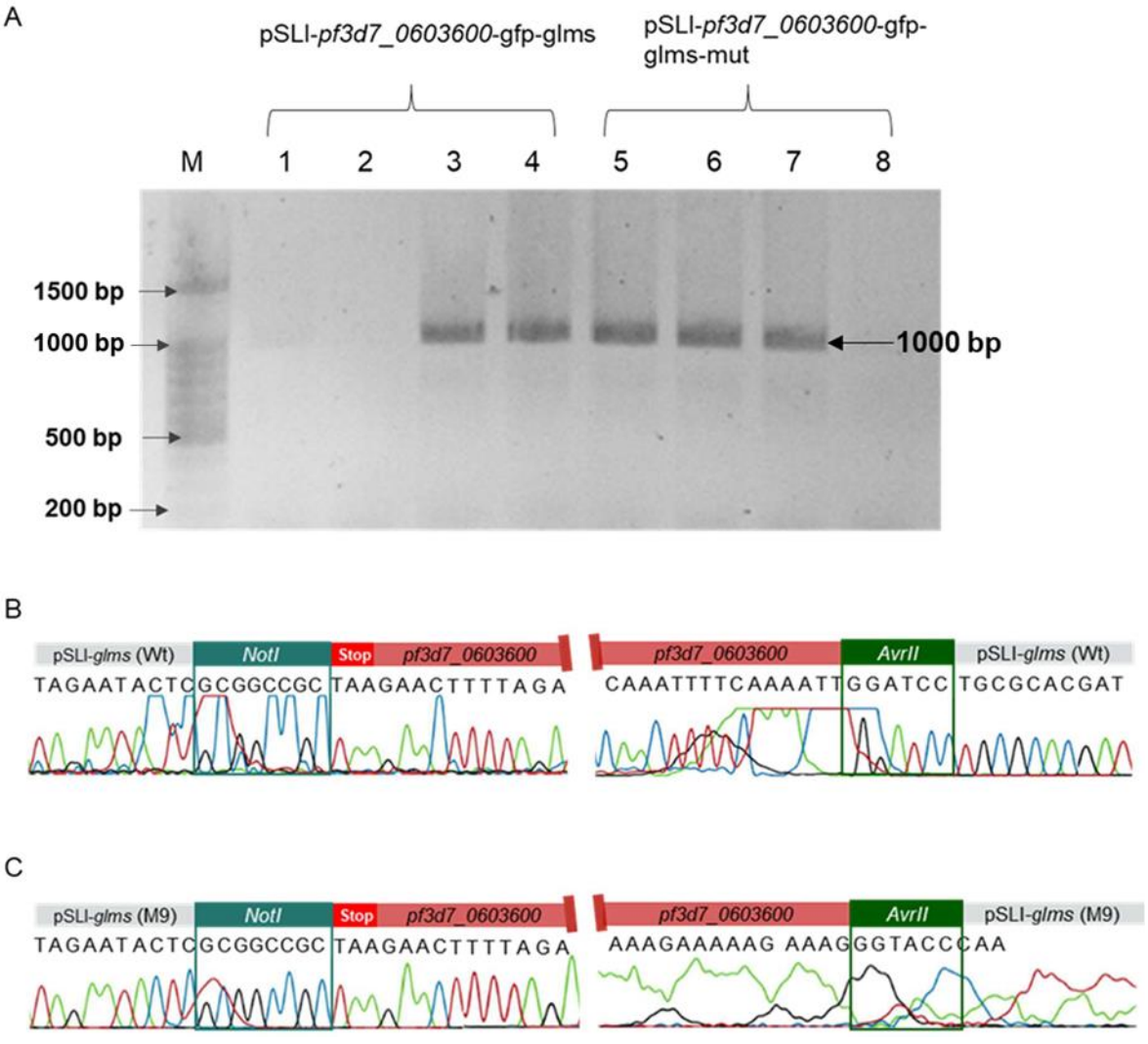

Figure 3-21. Production of pSLI-*pf3d7_0603600*-gfp-glms and pSLI-*pf3d7_0603600*-gfp-glms-mut plasmids. **A.** Colonies were screened for the presence of the plasmid using gene-specific primers, ASF and ASR. For each construct, four colonies were screened. For pSLI-*pf3d7_0603600*-gfp-glms two were positive and for pSLI-*pf3d7_0603600*-gfp-glms-mut three were positive, showing a band of approximately 950 bp which corresponded to the theoretical size of 969 bp. The reactions were observed on a 1.5 % agarose/TAE gel stained with EtBr. **B.** Colony three of pSLI-*pf3d7_0603600*-gfp-glms was sequenced the backbone, insert and both restriction sites were observed without any mutations. C. Colony five of pSLI-*pf3d7_0603600*-gfp-glms-mut was sequenced and the expected sequences were observed without any mutations present.

3.2.7 Production of a conditional knock-down system for Pf3D7_0603600

Once the plasmids were verified as being correct and without mutations, 100 µg of each plasmid was transfected into a majority ring NF54 culture. Time constants of pSLI-*pf3d7_0603600*-gfp-glms was 13.3 ms and pSLI-*pf3d7_0603600*-gfp-glms-mut was 18.7 ms, both of which are in the recommended range. NF54-epi(pSLI-*pf3d7_0603600*-gfp-glms) parasites were not observed 4 days after drug treatment and were observed again 17 days after drug pressure was removed. NF54-epi(pSLI-*pf3d7_0603600*-gfp-glms-mut) parasites were unobservable after 3 days and recovered after 35 days (Figure 3-22A). Episomal uptake was confirmed with a PCR using the SLI-F and GFP-R backbone specific primers. A band of approximately 1000 bp was observed which corresponded to the expected size of 969 bp(Figure 3-22B).

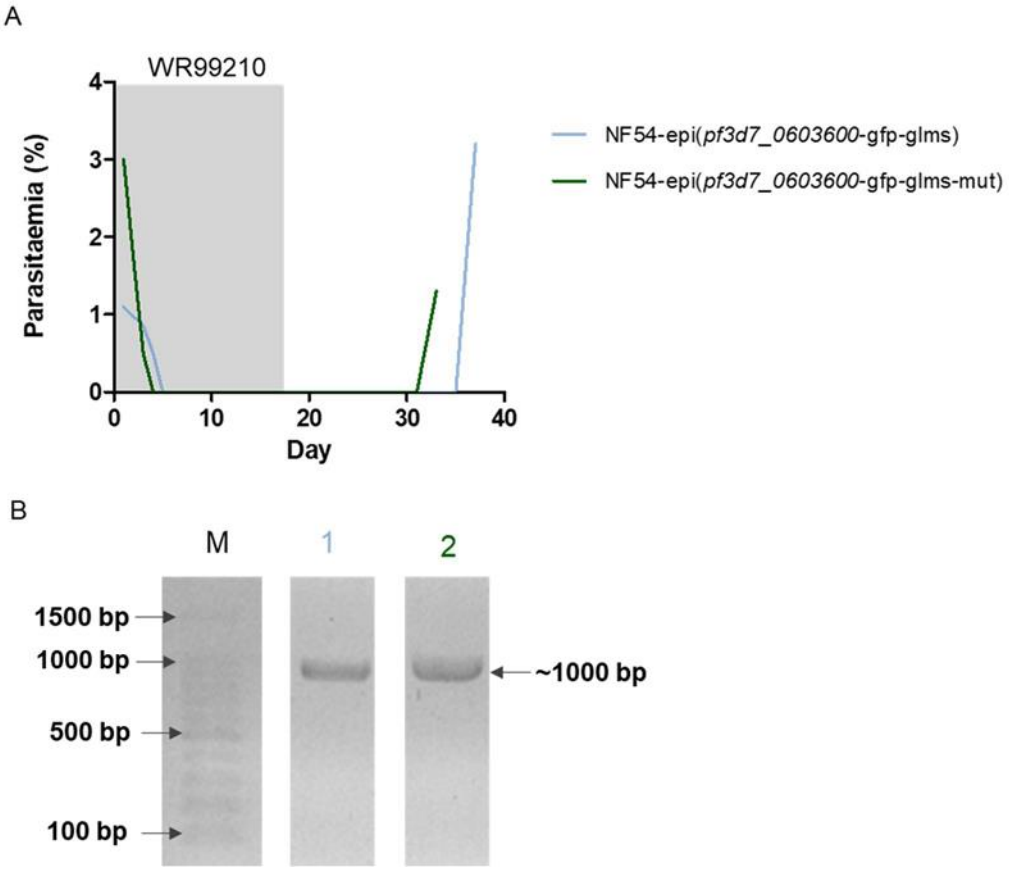

Figure 3-22. Production of NF54-epi(pSLI-*pf3d7_0603600*-gfp-*glms*) and NF54-epi(pSLI-*pf3d7_0603600*-gfp-*glms-mut*) lines. A. The time taken for parasitaemia to decrease and recover while on WR99210 treatment for episomal selection. The grey block indicates the amount of time the media was supplemented with WR99210. B. Lane M; A 100 bp molecular marker. Lane 1: A PCR with backbone specific primers (SLI-F and GFP-R) was done to confirm episomal uptake for pSLI-glms. A band of approximately 1000 bp was observed which corresponded to the theoretical size of 969 bp. Lane 2: the episomal uptake of pSLI-*Pf3D7_0603600*-gfp-*glms-mut* transfected parasites was confirmed with PCR. A band of approximately 1000 bp was observed. The reaction was visualised on a 1 % agarose/TAE gel poststained with EtBr.

Once episomal uptake was confirmed the parasites were selected for integration. The parasitaemia was not observable after 5 days of drug treatment and recovered after 25 days

from starting drug pressure (Figure 3-23A). A PCR to determine the presence of WT parasites was done and bands of approximately 1100 bp were observed in all three cultures. This corresponded to the expected size of 1211 bp. Integration was not observed in either the 5' or 3' region (Figure 3-23B) which does not correspond to the expected band sizes of the integrated lines. A band of 1215 bp was expected for the 5' PCR and a band of 1245 bp was expected for the 3' integration PCR. Thus, since there were bands present in the locus PCR and no bands in either of the integration PCRs, it can be concluded that no integrated parasites were generated for this line.

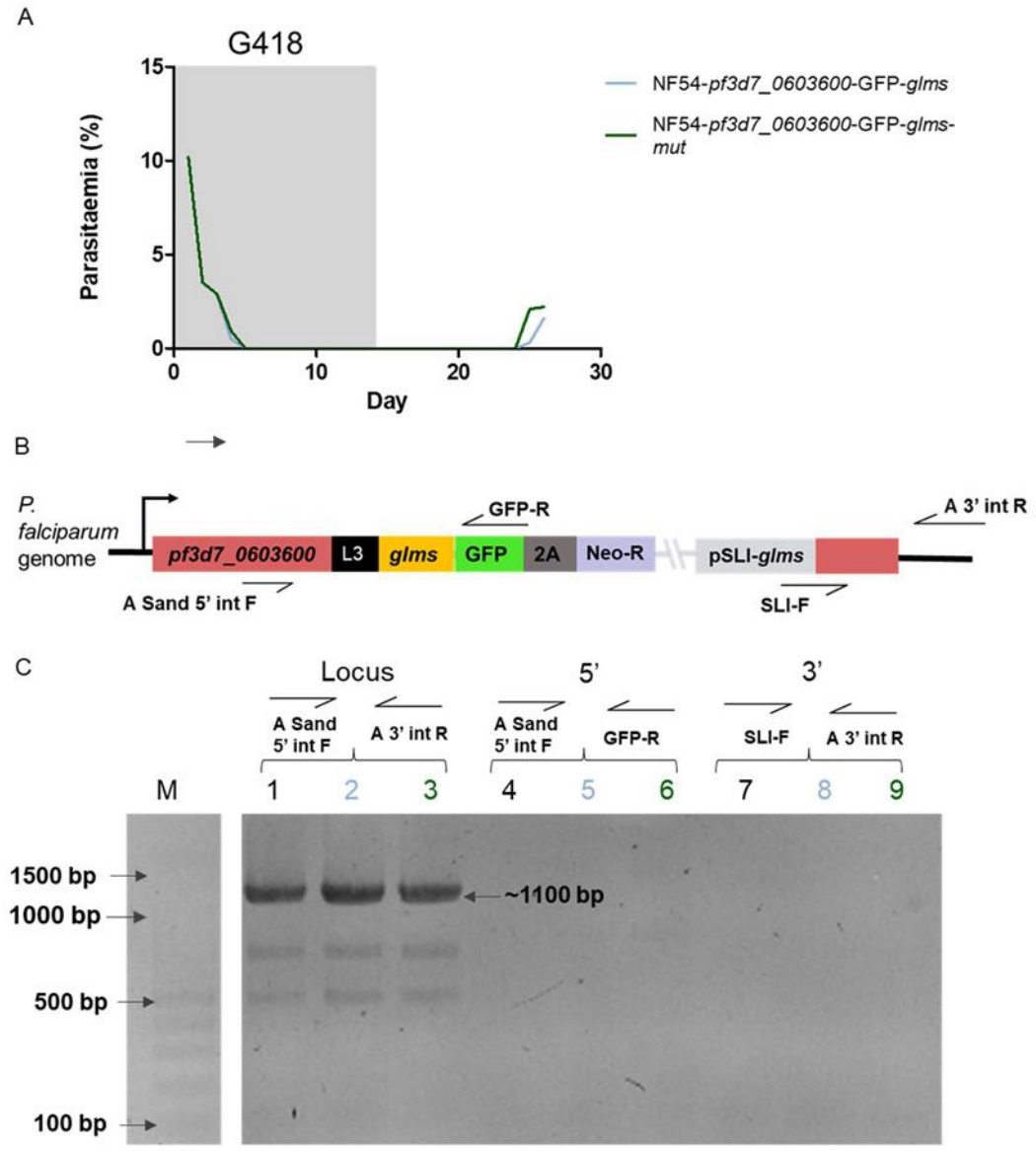

Figure 3-23. Selection for NF54-*pf3d7_0603600*-GFP-*glms* and NF54-*pf3d7_0603600*-GFP-*glms-mut* integrated lines. A. Growth of parasites during G418 drug selection. Parasites were not observed after 5 days of drug treatment and were observed again after 11 days of recovery. B. A schematic of the primers used to test for integration. C. Integration PCR of NF54-*PF3D7_0603600*-GFP-*glms* and NF54-*PF3D7_0603600*-GFP-*glms-mut*. Bands of approximately 1100 bp were observed which corresponded with the expected size of 1211 bp. No bands were observed in the 5' or 3' integration reactions. Thus, no integration was obtained for either of these lines. Thereactions were visualised on a 1.5 % agarose/TAE gel stained with EtBr.

Thus, Pf3D7_0603600 contains an ARID DBD and is not essential for asexual stages as validated by TGD.

4 Chapter 4: Discussion

P. falciparum is a complex organism that requires stringent gene expression regulation to effectively proliferate and differentiate. The parasite utilises many different methods of gene expression regulation on different levels such as epigenetic (36), transcriptional, post-transcriptional and post-translational. Recently studies on how epigenetics regulates the parasite genome has increased and it has been found to be highly specific (43). Transcriptional regulation occurs by both epigenetic mechanisms and various elements such as *cis*- and *trans*-acting elements (26). TF are *trans*-acting elements that bind to regulatory elements (*cis*-acting elements) to regulate transcription. The largest family of TF is the ApiAP2 family with 27 members which have been found to be master regulators of many life cycle stages in *P. falciparum* (66,71,112). This mechanism of master regulation is important in *P. falciparum* as it employs a just-in-time manner of gene expression likely caused by RNA pol pausing and an all-at-once manner of transcription (113). The parasite likely "prepares" all the genes for needed transcripts by assembling the PIC before initiating a life cycle stage change and employing the all-at-once method of transcription.

Even though the ApiAP2 family acts as master regulators and are thought to be the most important TF for life cycle progression in the parasite, during transcriptional studies, it has been found that two TF, PfMyb1 and Pf3D7_0603600 may be important during early-stage differentiation (72). Understanding the functions of these TF may be useful to improve our understanding of differentiation and improve transmission-blocking efforts. There is a very high gene to TF ratio in *P. falciparum* (114) suggesting that TF are either master regulators or other gene regulation systems are in place or both. This suggests that there are multiple layers of gene expression regulation that must be further studied to understand the biology of this malaria-causing parasite.

PfMyb1 was confirmed as a TF because it contains a homeobox and myb domain. The myb domain binds DNA and SANT proteins which is likely that Myb1 not only binds DNA directly but is part of a larger complex. PfMyb1 acts as either an enhancer or repressor in the PIC (113) and the through MRE either increase or decrease expression of certain genes (57).

pfmyb1 is likely essential during the asexual stages due to the protein functioning as either an activator or repressor as well as the lncRNA regulating gene expression. The genes that Myb1 regulate cannot be specifically grouped to a single life stage or specific function but is important for both the proliferative cell cycle, DNA synthesis and metabolism.

The increased expression of *pfmyb1* during the early gametocyte stages may be explained by the role of *pfmyb1* in the PIC as either an enhancer or repressor which would indicate an

increase in transcription. The increased expression of *pfmyb1* in the late stages may be due to the increased expression in genes for mobility required by the gamete stages as well as the ability of female gametocytes to store transcripts for later stages. The decrease in expression during the gametocyte stages can be attributed to the decrease in DNA synthesis, transcription levels and overall metabolism.

In this study, *pfmyb1* was confirmed to be essential in the asexual stages by TGD. Since no recombinant line could be produced using TGD, due to *pfmyb1* essentiality, a conditional knock-down system was required to study the role of *pfmyb1* during the gametocyte stages.

It was found that the changes made to the *pfmyb1* locus by the conditional knock-down system were well tolerated by the parasite as no growth or morphological changes were observed. This likely means that the gene fragment recombined in the correct place (3') and orientation to allow expression of the recombinant *pfmyb1* gene. The success of the recombination was further proven by the observation of the reporter gene (GFP) in the nucleus of the parasite, as confirmed by Hoechst staining. PfMyb1 was localised to the nucleus of the parasite. This is the first time PfMyb1 has been localised in the parasite and because it is a TF it is expected to localise in the nucleus.

Pf3D7_0603600 has been annotated as a putative TF likely due to its' ARID DBD (72). This DBD is found in many eukaryotes and have multiple functions (79). Thus, it is difficult to determine the function of Pf3D7_0603600 other than it binds DNA.

It has previously been found to be dispensable in asexual stages (76) which was confirmed with TGD. Even though a "clean" line could not be produced, i.e., there were still WT parasites present in the culture, the integration PCR did show integration of the recombinant gene for asexual parasites meaning that *pf3d7_0603600* is not essential for the asexual stages.

The first two objectives, to produce a genetically recombinant line to allow targeted gene disruption and conditional knock down of *pfmyb1*, was achieved fully. Objective 3 was partially achieved as the line was not fully integrated and still contained WT parasites and objective 4 was not achieved as no integration at the *pf3d7_0603600* locus was observed. My hypothesis, PfMyb1 has functional effects during proliferation and differentiation while PF3D7_0603600 has functional effects during differentiation, was partially accepted in the sense that PfMyb1 does have functional effects during proliferation, due to its essentiality and Pf3D7_0603600 does not have functional effects during proliferation due to its dispensability. However, no data was produced for differentiation so no conclusions could be made. From this, it can be deduced that PfMyb1 responded well to the SLI system as recombinant parasites were obtained quickly and without WT parasites present however, Pf3D7_0603600 did not respond well as neither the TGD nor conditional knock-down lines produced "clean" recombinant lines.

This may be due to the differences in gene lengths (2195 bp vs 8809 bp). Another possibility is that recombination occurred in the wrong locus.

Shortcomings of the study include non-quantitative evidence for GFP expression in NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut*. This could be done using a different antibody for GFP for the Western blot or through qPCR. A "clean" line could not be obtained for NF54-Li*pf3d7_0603600*-gfp which means that it could not be used to determine the functional effects of Pf3D7_0603600 during differentiation. No integration was observed for NF54-*pf3d7_0603600*-GFP-*glms* or NF54-*pf3d7_0603600*-GFP-*glms-mut* thus, either parasites which have taken up the plasmid episomally would have to be put on G418 drug pressure again to try and obtain integration or if that is unsuccessful a different system will have to be used.

Future prospects include using the recombinant lines generated in this study to further evaluate the functional effects of PfMyb1 during differentiation through the use of the NF54-*pfmyb1*-GFP-*glms* and NF54-*pfmyb1*-GFP-*glms-mut* lines. Using the TGD line for *pf3d7_0603600,* this putative TF can be localised and a global gene study to see which genes are affected by p*f3d7_0603600* removal during proliferation and differentiation would show which genes are regulated by *pf3d7_0603600.*

5 Conclusion

Malaria is caused by a protozoan with a complex life cycle that requires stringent gene expression regulation. The parasite regulates its gene expression in a myriad of ways including TF. Two TF, PfMyb1 and PF3D7_0603600, were found to be upregulated in the early stages of gametocytogenesis and are thought to be important for differentiation. This was studied by generating recombinant lines for both genes. For Pf3D7_0603600 partial integration of a knockout line was achieved. For PfMyb1, no integration was observed in the knockout line showing that it is essential in asexual stages. A conditional knock down line was then generated. Integration was observed and GFP was visualised in the recombinant parasites using confocal microscopy. The line was compared to WT parasites to determine whether the genetic recombination has disturbed any of the general housekeeping functions of the parasite. No significant difference was observed in the growth rates between NF54, NF54-*Pfmyb1*-GFP-*glms* and NF54-*Pfmyb1*-GFP-*glms-mut* in both the asexual and sexual stages of the parasites life cycle. Morphology was also compared, and no visible differences were observed. These conditional knock down lines can be used to further study the effect of PfMyb1 during differentiation by parasite growth and perturbed gene expression during the sexual stages to determine the genes whose expression is affected by PfMyb1.

6 References

1. (2020) World malaria report 2020: 20 years of global progress and challenges. World Health Organisation, Geneva
2. White, N. J., Pukrittayakamee, S., Hien, T. T., Faiz, M. A., Mokuolu, O. A., and Dondorp, A. M. (2014) Malaria. *The Lancet* **383**, 723-735
3. Ta, T. H., Hisam, S., Lanza, M., Jiram, A. I., Ismail, N., and Rubio, J. M. (2014) First case of a naturally acquired human infection with *Plasmodium cynomolgi*. *Malaria Journal* **13**, 68
4. Krotoski, W. A., Krotoski, D., Garnham, P., Bray, R., Killick-Kendrick, R., Draper, C., Targett, G., and Guy, M. (1980) Relapses in primate malaria: discovery of two populations of exoerythrocytic stages. Preliminary note. *British Medical Journal* **280**, 153
5. Adams, J. H., and Mueller, I. (2017) The biology of *Plasmodium vivax*. *Cold Spring Harb Perspectives in Medicine*, a025585
6. Mueller, I., Zimmerman, P. A., and Reeder, J. C. (2007) *Plasmodium malariae* and *Plasmodium ovale*-the 'bashful'malaria parasites. *Trends in Parasitology* **23**, 278-283
7. Cox-Singh, J., and Singh, B. (2008) Knowlesi malaria: newly emergent and of public health importance? *Trends in Parasitology* **24**, 406-410
8. Faye, F. B., Spiegel, A., Tall, A., Sokhna, C., Fontenille, D., Rogier, C., and Trape, J.-F. (2002) Diagnostic criteria and risk factors for *Plasmodium ovale* malaria. *The Journal of Infectious Diseases* **186**, 690-695
9. WHO. (2017) World Malaria Report 2018 Geneva
10. Idro, R., Marsh, K., John, C. C., and Newton, C. R. (2010) Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatric Research* **68**, 267-274
11. Hanson, J., Lee, S. J., Hossain, M. A., Anstey, N. M., Charunwatthana, P., Maude, R. J., Kingston, H. W., Mishra, S. K., Mohanty, S., and Plewes, K. (2015) Microvascular obstruction and endothelial activation are independently associated with the clinical manifestations of severe falciparum malaria in adults: an observational study. *BMC Medicine* **13**, 122
12. WHO. (2018) Global report on insecticide resistance in malaria vectors: 2010-2016 (Organisation, W. H. ed., Geneva
13. Gosling, R., and von Seidlein, L. (2016) The future of the RTS, S/AS01 malaria vaccine: an alternative development plan. *PLoS Medicine* **13**, e1001994
14. Lucille Blumberg, J. F., Natalie Mayet, Jaishree, and Raman, D. I. S. U. (2017) SOUTH AFRICAN GUIDELINES FOR THE PREVENTION OF MALARIA. (Health, D. o. ed.
15. White, N. (1999) Antimalarial drug resistance and combination chemotherapy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **354**, 739-749
16. Nosten, F., and White, N. J. (2007) Artemisinin-based combination treatment of falciparum malaria. *The American Journal of Tropical Medicine and Hygiene* **77**, 181-192
17. Abba, K., Deeks, J. J., Olliaro, P. L., Naing, C. M., Jackson, S. M., Takwoingi, Y., Donegan, S., and Garner, P. (2011) Rapid diagnostic tests for diagnosing uncomplicated *P. falciparum* malaria in endemic countries. *Cochrane Database of Systematic Reviews*
18. Baker, D. A. (2010) Malaria gametocytogenesis. *Molecular and Biochemical Parasitology* **172**,57-65
19. Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J., and DeRisi, J. L. (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* **1**, e5
20. Weiner, A., Dahan-Pasternak, N., Shimoni, E., Shinder, V., von Huth, P., Elbaum, M., and Dzikowski, R. (2011) 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite *Plasmodium falciparum*. *Cellular Microbiology* **13**, 967-977
21. Lu, X. M., Batugedara, G., Lee, M., Prudhomme, J., Bunnik, E. M., and Le Roch, K. G. (2017) Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Research* **45**, 7825-7840
22. Ay, F., Bunnik, E. M., Varoquaux, N., Bol, S. M., Prudhomme, J., Vert, J.-P., Noble, W. S., and Le Roch, K. G. (2014) Three-dimensional modeling of the *P. falciparum* genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Research* **24**, 974-988
23. Kensche, P. R., Hoeijmakers, W. A. M., Toenhake, C. G., Bras, M., Chappell, L., Berriman, M., and Bartfai, R. (2016) The nucleosome landscape of *Plasmodium falciparum* reveals

chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Research* **44**,2110-2124
24. Read, D. F., Cook, K., Lu, Y. Y., Le Roch, K. G., and Noble, W. S. (2019) Predicting gene expression in the human malaria parasite *Plasmodium falciparum* using histone modification, nucleosome positioning, and 3D localization features. *PLoS Computational Biology* **15**, e1007329
25. Santos, J. M., Josling, G., Ross, P., Joshi, P., Orchard, L., Campbell, T., Schieler, A., Cristea, I. M., and Llinas, M. (2017) Red blood cell invasion by the malaria parasite is coordinated by the PfAP2-I transcription factor. *Cell Host & Microbe* **21**, 731-741. e710
26. Howick, V. M., Russell, A. J., Andrews, T., Heaton, H., Reid, A. J., Natarajan, K., Butungi, H., Metcalf, T., Verzier, L. H., and Rayner, J. C. (2019) The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* **365**
27. Ruecker, A., Mathias, D., Straschil, U., Churcher, T., Dinglasan, R., Leroy, D., Sinden, R., and Delves, M. (2014) A male and female gametocyte functional viability assay to identify biologically relevant malaria transmission-blocking drugs. *Antimicrobial Agents and Chemotherapy* **58**, 7292-7302
28. Kuehn, A., and Pradel, G. (2010) The coming-out of malaria gametocytes. *BioMed Research International* **2010**
29. Farfour, E., Charlotte, F., Settegrana, C., Miyara, M., and Buffet, P. (2012) The extravascular compartment of the bone marrow: a niche for *Plasmodium falciparum* gametocyte maturation? *Malaria Journal* **11**, 285
30. Silvestrini, F., Alano, P., and Williams, J. (2000) Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*. *Parasitology* **121**, 465-471
31. Coleman, B. I., Skillman, K. M., Jiang, R. H., Childs, L. M., Altenhofen, L. M., Ganter, M., Leung, Y., Goldowitz, I., Kafsack, B. F., and Marti, M. (2014) A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host &Microbe* **16**, 177-186
32. Brancucci, N. M., Bertschi, N. L., Zhu, L., Niederwieser, I., Chin, W. H., Wampfler, R., Freymond, C., Rottmann, M., Felger, I., and Bozdech, Z. (2014) Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host & Microbe* **16**, 165-176
33. Filarsky, M., Fraschka, S. A., Niederwieser, I., Brancucci, N. M., Carrington, E., Carrio, E., Moes, S., Jenoe, P., Bartfai, R., and Voss, T. S. (2018) GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259-1263
34. Connacher, J., Josling, G. A., Orchard, L. M., Reader, J., Llinas, M., and Birkholtz, L.-M. (2021) H3K36 methylation reprograms gene expression to drive early gametocyte development in *Plasmodium falciparum*. *Epigenetics & Chromatin* **14**, 1-15
35. Yuda, M., Kaneko, I., Iwanaga, S., Murata, Y., and Kato, T. (2020) Female-specific gene regulation in malaria parasites by an AP2-family transcription factor. *Molecular Microbiology* **113**, 40-51
36. Coetzee, N., von Gruning, H., Opperman, D., van der Watt, M., Reader, J., and Birkholtz, L.-M. (2020) Epigenetic inhibitors target multiple stages of *Plasmodium falciparum* parasites. *Scientific Reports* **10**, 1-11
37. Kaneko, I., Iwanaga, S., Kato, T., Kobayashi, I., and Yuda, M. (2015) Genome-wide identification of the target genes of AP2-O, a *Plasmodium* AP2-family transcription factor. *PLoS Pathogens* **11**, e1004905
38. Khan, S. M., Franke-Fayard, B., Mair, G. R., Lasonder, E., Janse, C. J., Mann, M., and Waters, A. P. (2005) Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell* **121**, 675-687
39. Mair, G. R., Lasonder, E., Garver, L. S., Franke-Fayard, B. M., Carret, C. K., Wiegant, J. C., Dirks, R. W., Dimopoulos, G., Janse, C. J., and Waters, A. P. (2010) Universal features of post-transcriptional gene regulation are critical for *Plasmodium* zygote development. *PLoS Pathogens* **6**, e1000767
40. Adjalley, S. H., Chabbert, C. D., Klaus, B., Pelechano, V., and Steinmetz, L. M. (2016) Landscape and dynamics of transcription initiation in the malaria parasite *Plasmodium falciparum*. *Cell Reports* **14**, 2463-2475
41. Miao, J., Li, J., Fan, Q., Li, X., Li, X., and Cui, L. (2010) The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite *Plasmodium falciparum*. *Journal of Cell Science* **123**, 1039-1049

42. van Biljon, R., Niemand, J., van Wyk, R., Clark, K., Verlinden, B., Abrie, C., von Gruning, H., Smidt, W., Smit, A., and Reader, J. (2018) Inducing controlled cell cycle arrest and re-entry during asexual proliferation of *Plasmodium falciparum* malaria parasites. *Scientific reports* **8**, 16581
43. Fraschka, S. A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X. Y., Brancucci, N. M., Mohring, F., Mushunje, A. T., Huang, X., and Christensen, P. R. (2018) Comparative heterochromatin profiling reveals conserved and unique epigenome signatures linked to adaptation and development of malaria parasites. *Cell Host & Microbe* **23**, 407-420. e408
44. Ruiz, J. L., Tena, J. J., Bancells, C., Cortes, A., Gomez-Skarmeta, J. L., and Gomez-Dfaz, E. (2018) Characterization of the accessible genome in the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Research* **46**, 9414-9431
45. Josling, G. A., Russell, T. J., Venezia, J., Orchard, L., van Biljon, R., Painter, H. J., and Llinas, M. (2020) Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nature Communications* **11**, 1-13
46. Doerig, C., Rayner, J. C., Scherf, A., and Tobin, A. B. (2015) Post-translational protein modifications in malaria parasites. *Nature Reviews Microbiology* **13**, 160-172
47. Solyakov, L., Halbert, J., Alam, M. M., Semblat, J.-P., Dorin-Semblat, D., Reininger, L., Bottrill, A. R., Mistry, S., Abdi, A., and Fennell, C. (2011) Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. *Nature Communications* **2**, 1-12
48. Bannister, A. J., and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell research* **21**, 381-395
49. Sorber, K., Dimon, M. T., and DeRisi, J. L. (2011) RNA-Seq analysis of splicing in *Plasmodium falciparum* uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Research* **39**, 3820-3835
50. Miao, J., Fan, Q., Parker, D., Li, X., Li, J., and Cui, L. (2013) Puf mediates translation repression of transmission-blocking vaccine candidates in malaria parasites. *PLoS Pathogens* **9**, e1003268
51. Shock, J. L., Fischer, K. F., and DeRisi, J. L. (2007) Whole-genome analysis of mRNA decay in *Plasmodium falciparum* reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. *Genome Biology* **8**, 1-12
52. Vembar, S. S., Droll, D., and Scherf, A. (2016) Translational regulation in blood stages of the malaria parasite *Plasmodium spp*.: systems-wide studies pave the way. *Wiley Interdisciplinary Reviews: RNA* **7**, 772-792
53. Coulson, R. M., Hall, N., and Ouzounis, C. A. (2004) Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Research* **14**, 1548-1554
54. Gopalakrishnan, A. M., Nyindodo, L. A., Fergus, M. R., and Lopez-Estrano, C. (2009) *Plasmodium falciparum*: preinitiation complex occupancy of active and inactive promoters during erythrocytic stage. *Experimental parasitology* **121**, 46-54
55. Levine, M., and Tjian, R. (2003) Transcription regulation and animal diversity. *Nature* **424**, 147
56. Templeton, T. J., Iyer, L. M., Anantharaman, V., Enomoto, S., Abrahante, J. E., Subramanian, G., Hoffman, S. L., Abrahamsen, M. S., and Aravind, L. (2004) Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Research* **14**, 1686-1695
57. Gissot, M., Briquet, S., Refour, P., Boschet, C., and Vaquero, C. (2005) PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation. *Journal of Molecular Biology* **346**, 29-42
58. Kafsack, B. F., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., Williams, A. E., Drought, L. G., Kwiatkowski, D. P., and Baker, D. A. (2014) A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248
59. Chou, E. S., Abidi, S. Z., Teye, M., Leliwa-Sytek, A., Rask, T. S., Cobbold, S. A., Tonkin-Hill, G. Q., Subramaniam, K. S., Sexton, A. E., and Creek, D. J. (2018) A high parasite density environment induces transcriptional changes and cell death in *Plasmodium falciparum* blood stages. *The FEBS Journal* **285**, 848-870
60. Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M. (1998) A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal* **17**, 5484-5496
61. Riechmann, J. L., and Meyerowitz, E. M. (1998) The AP2/EREBP family of plant transcription factors. *Biological Chemistry* **379**, 633-646

62. Balaji, S., Babu, M. M., Iyer, L. M., and Aravind, L. (2005) Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Research* **33**, 3994-4006
63. Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J., and DeRisi, J. L. (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biology* **1**, e5
64. Le Roch, K. G., Zhou, Y., Blair, P. L., Grainger, M., Moch, J. K., Haynes, J. D., De la Vega, P., Holder, A. A., Batalov, S., and Carucci, D. J. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503-1508
65. Yuda, M., Iwanaga, S., Shigenobu, S., Mair, G. R., Janse, C. J., Waters, A. P., Kato, T., and Kaneko, I. (2009) Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Molecular Microbiology* **71**, 1402-1414
66. Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., and Kaneko, I. (2010) Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Molecular Microbiology* **75**, 854-863
67. Balu, B., Chauhan, C., Maher, S. P., Shoue, D. A., Kissinger, J. C., Fraser, M. J., and Adams, J. H. (2009) *piggyBac* is an effective tool for functional analysis of the *Plasmodium falciparum* genome. *BMC Microbiology* **9**, 83
68. Maier, A. G., Rug, M., O'Neill, M. T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., and Coppel, R. L. (2008) Exported proteins required for virulence and rigidity of *Plasmodium falciparum*-infected human erythrocytes. *Cell* **134**, 48-61
69. Martins, R. M., Macpherson, C. R., Claes, A., Scheidig-Benatar, C., Sakamoto, H., Yam, X. Y., Preiser, P., Goel, S., Wahlgren, M., and Sismeiro, O. (2017) An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite *Plasmodium falciparum*. *Scientific Reports* **7**, 1-10
70. Singh, S. (2019) Characterizing the functional role of the transcriptional regulator PfAP2-G2 in *Plasmodium falciparum* gametocytogenesis. Dissertation.
71. Xu, Y., Qiao, D., Wen, Y., Bi, Y., Chen, Y., Huang, Z., Cui, L., Guo, J., and Cao, Y. (2021) PfAP2-G2 Is Associated to Production and Maturation of Gametocytes in *Plasmodium falciparum* via Regulating the Expression of PfMDV-1. *Frontiers in Microbiology* **11**, 3546
72. van Biljon, R., van Wyk, R., Painter, H. J., Orchard, L., Reader, J., Niemand, J., Llinas, M., and Birkholtz, L. (2018) Intricate hierarchical transcriptional control regulates *Plasmodium falciparum* sexual differentiation. *BMC Genomics* **20**, 920
73. Bancells, C., Llora-Batlle, O., Poran, A., Notzel, C., Rovira-Graells, N., Elemento, O., Kafsack, B. F., and Cortes, A. (2019) Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*. *Nature Microbiology* **4**, 144-154
74. Carrington, E., Cooijmans, R. H. M., Keller, D., Toenhake, C. G., Bartfai, R., and Voss, T. S. (2021) The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in the malaria parasite *Plasmodium falciparum*. *Iscience* **24**, 102444
75. Boschet, C., Gissot, M., Briquet, S., Hamid, Z., Claudel-Renard, C., and Vaquero, C. (2004) Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones. *Molecular and Biochemical Parasitology* **138**, 159
76. Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D., and Mayho, M. (2018) Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360**, 6388
77. Iwahara, J., and Clubb, R. T. (1999) Solution structure of the DNA binding domain from Dead ringer, a sequence-specific AT-rich interaction domain (ARID). *The EMBO journal* **18**, 6084-6094
78. Whitson, R. H., Huang, T., and Itakura, K. (1999) The novel Mrf-2 DNA-binding domain recognizes a five-base core sequence through major and minor-groove contacts. *Biochemical and biophysical research communications* **258**, 326-331
79. Wilsker, D., Patsialou, A., Dallas, P. B., and Moran, E. (2002) ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. *Heart* **28**, 30
80. Gil Carvalho, T., and Menard, R. (2005) Manipulating the *Plasmodium* genome. *Current Issues in Molecular Biology* **7**, 39-56
81. Cowman, A. F., and Crabb, B. S. (2005) Genetic manipulation of Plasmodium falciparum. *Molecular Approaches to Malaria*, 50-67
82. Crabb, B. S., and Cowman, A. F. (1996) Characterization of promoters and stable transfection by homologous and nonhomologous recombination in *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences* **93**, 7289-7294

83. O'Donnell, R. A., Preiser, P. R., Williamson, D. H., Moore, P. W., Cowman, A. F., and Crabb, B. S. (2001) An alteration in concatameric structure is associated with efficient segregation of plasmids in transfected *Plasmodium falciparum* parasites. *Nucleic Acids Research* **29**, 716-724
84. Waterkeyn, J., Crabb, B., and Cowman, A. (1999) Transfection of the human malaria parasite *Plasmodium falciparum*. *International Journal for Parasitology* **29**, 945-955
85. Crabb, B. S., Rug, M., Gilberger, T.-W., Thompson, J. K., Triglia, T., Maier, A. G., and Cowman, A. F. (2004) Transfection of the human malaria parasite *Plasmodium falciparum*. in *Parasite Genomics Protocols*, Springer. pp 263-276
86. Birnbaum, J., Flemming, S., Reichard, N., Soares, A. B., Mesen-Ramfrez, P., Jonscher, E., Bergmann, B., and Spielmann, T. (2017) Selection linked integration (SLI) for endogenous gene tagging and knock sideways in *Plasmodium falciparum* parasites. *Nat Methods* **14**, 450-456
87. Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., and Duquenne, L. (2009) InterPro: the integrative protein signature database. *Nucleic Acids Research* **37**, D211-D215
88. Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., Brown, S. D., Chang, H.-Y., El-Gebali, S., and Fraser, M. I. (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research* **47**, D351-D360
89. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., and Bordoli, L. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296-W303
90. Kieleczawa, J. (2006) Fundamentals of sequencing of difficult templates—an overview. *Journal of Biomolecular Techniques: JBT* **17**, 207
91. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press
92. Whitaker, J. R., and Granum, P. E. (1980) An absolute method for protein determination based on difference in absorbance at 235 and 280 nm. *Analytical Biochemistry* **109**, 156-159
93. Wahlgren, I. L. m. K. M. H. P. A. S. M. (2008) Methods in Malaria Research. (Kirsten Moll, I. L. m., Hedvig Perlmann, Artur Scherf, Mats Wahlgren. ed., 5 Ed., MR4/ATCC Manassas, VA
94. Singh, B., Bobogare, A., Cox-Singh, J., Snounou, G., Abdullah, M. S., and Rahman, H. A. (1999) A genus-and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies. *American Journal of Tropical Medicine and Hygiene* **60**, 687-92
95. Mandel, M., and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53**, 159-162
96. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467
97. Allen, R. J. & Kirk, K. 2010. Plasmodium falciparum culture: the benefits of shaking. *Molecular and Biochemical Parasitology* **169**, 63-5.
98. Trager, W., and Jensen, J. B. (1976) Human malaria parasites in continuous culture. *Science* **193**, 673-675
99. Verlinden, B. K., Niemand, J., Snyman, J., Sharma, S. K., Beattie, R. J., Woster, P. M., and Birkholtz, L.-M. (2011) Discovery of novel alkylated (bis)urea and (bis)thiourea polyamine analogues with potent antimalarial activities. *Journal of Medicinal Chemistry* **54**, 6624-6633
100. Warhurst, D., and Williams, J. (1996) ACP Broadsheet no 148. July 1996. Laboratory diagnosis of malaria. *Journal of Clinical Pathology* **49**, 533
101. Wu, Y., Sifri, C. D., Lei, H.-H., Su, X.-z., and Wellems, T. E. (1995) Transfection of *Plasmodium falciparum* within human red blood cells. *Proceedings of the National Academy of Sciences* **92**, 973-977
102. Chaianantakul, N., Sirawaraporn, R., and Sirawaraporn, W. (2013) Insights into the role of the junctional region of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Malaria Journal* **12**, 1-13
103. Kotra, L. P., Haddad, J., and Mobashery, S. (2000) Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. *Antimicrobial Agents and Chemotherapy* **44**, 3249-3256
104. Mamoun, C. B., Gluzman, I. Y., Goyard, S., Beverley, S. M., and Goldberg, D. E. (1999) A set of independent selectable markers for transfection of the human malaria parasite *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences* **96**, 8716-8720
105. Fernandez, V. (2004) Sorbitol-synchronization of *Plasmodium falciparum*-infected

erythrocytes. *Methods in Malaria Research, 4th ed. MR4/ATCC, Manassas, VA*, 24

106. Reader, J., Botha, M., Theron, A., Lauterbach, S. B., Rossouw, C., Engelbrecht, D., Wepener, M., Smit, A., Leroy, D., and Mancama, D. (2015) Nowhere to hide: interrogating different

metabolic parameters of *Plasmodium falciparum* gametocytes in a transmission blocking drug discovery pipeline towards malaria elimination. *Malaria Journal* **14**, 1-17
107. Ponnudurai, T., Lensen, A., Leeuwenberg, A. D., and Meuwissen, J. T. (1982) Cultivation of fertile *Plasmodium falciparum* gametocytes in semi-automated systems. 1. Static cultures. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **76**, 812-818
108. Briquet, S., Ourimi, A., Pionneau, C., Bernardes, J., Carbone, A., Chardonnet, S., and Vaquero, C. (2018) Identification of *Plasmodium falciparum* nuclear proteins by mass spectrometry and proposed protein annotation. *PLoS One* **13**, e0205596
109. Glauert, A. M., Dingle, J., and Lucy, J. (1962) Action of saponin on biological cell membranes. *Nature* **196**, 953
110. Mannervik, M. (1999) Target genes of homeodomain proteins. *BioEssays* **21**, 267-270
111. Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic acids research* **16**, 6127-6145
112. Lindner, S. E., De Silva, E. K., Keck, J. L., and Llinas, M. (2010) Structural determinants of DNA binding by a *P. falciparum* ApiAP2 transcriptional regulator. *Journal of Molecular Biology***395**, 558-567
113. Luse, D. S. (2014) The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled? *Transcription* **5**, e27050
114. Bischoff, E., and Vaquero, C. (2010) *In silico* and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of *Plasmodium falciparum*. *BMC genomics* **11**, 1-20

7 Supplementary

7.1 Plasmids used

Plasmid name	Description	Cloning sites	Cloning primers	Purpose	Source
PfMyb1-TGD	*ΔPfmyb1*-GFP	*NotI*, *MluI*	MTF MTR MSR	Disruption	pSLI-TGD(86)
PfMyb1-glms	*Pfmyb1*-GFP-*glms*	*NotI*, *AvrII*	MSF MSR	Knockdown	pSLI-*glms* (WT)
PfMyb1-mut	*Pfmyb1*-GFP-*glms*(M9)	*NotI*, *AvrII*	MSF MSR	Knockdown control	pSLI-*glms* (M9)
ARID-TGD	*ΔPf3D7_0603600* -GFP	*NotI*, *MluI*	ATF ATR ASR	Disruption	pSLI-TGD(86)
ARID-glms	*Pf3D7_0603600* -GFP-*glms*	*NotI*, *AvrII*	ASF ASR	Knockdown	pSLI-*glms* (WT)
ARID-mut	*Pf3D7_0603600* -GFP-*glms*(M9)	*NotI*, *AvrII*	ASF ASR	Knockdown control	pSLI-*glms* (M9)

7.2 Primers used

Primer name	Purpose	Position	Sequence
MTF	Cloning	Chr13:F661244-R660779	CGATGCGGCCGCTAAGAATATGAGTGTTTATGT
MTR	Cloning		CGATACGCGTTTGCTTTGGCTCACTTTG
MSF	Cloning	Chr13:F660991-R660018	GCATGCGGCCGCTAAATATTAGATTCACAATTAATG
MSR	Cloning		CGATCCTAGGTTGTTCCATGTTATTTGTTGCAC
ATF	Cloning	Chr6:F148886-R144897	CGATGCGGCCGCTAAGATATTCCTTTAATGGTTCC
ATR	Cloning		CGGCATACGCGTAGCACCACATATATTATTTTC
ASF	Cloning	Chr6:F142376-R141421	GCATGCGGCCGCTAAGAACTTTTAGATAATACG
ASR	Cloning		CGATCCTAGGTTTAATCTTTTCAACCTTTACCAACTTTG
SLI-F	Sequencing, episomal uptake, integration		AGCGGATAACAATTTCACACAGGA
GFP-R	Sequencing, episomal uptake, integration		ACAAGAATTGGGACAACTCCAGTGA
PfMyb1-5' int F	Integration	Chr13:F661550 - R659846	GAAGTTTGATTGTAGTACGAATTGGTAG
PfMyb1-3' int R	Integration		GTGATACATAATATCCAATATGCTAAC
A5' int F	Integration	Chr6:F148991 - R148453	GTACATATTTCCACGGGGTTAGG
ATGD 3'R	Integration		CCATTCAAATTATTTGTCAGAGCACC
ASand int F	Integration	Chr6:F149176 - R141259	CCTTAACATTCCTGTTGCTAAAAAGTG
A 3' int R	Integration		CCATTTCATTTTTTTATATTCTCCCC